SUPPORTING INFORMATION

MICROSCLERODERMINS FROM TERRESTRIAL MYXOBACTERIA:

AN INTRIGUING BIOSYNTHESIS LIKELY CONNECTED TO A SPONGE SYMBIONT

Thomas Hoffmann, Stefan Müller, Suvd Nadmid, Ronald Garcia and Rolf Müller*

Department of Microbial Natural Products Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS),
Helmholtz Centre for Infection Research (HZI) and Pharmaceutical Biotechnology, Saarland University, Campus C23, 66123 Saarbrücken (Germany)

rom@helmholtz-hzi.de

Table of Contents

Structure elucidation

NMR data of Microsclerodermin D

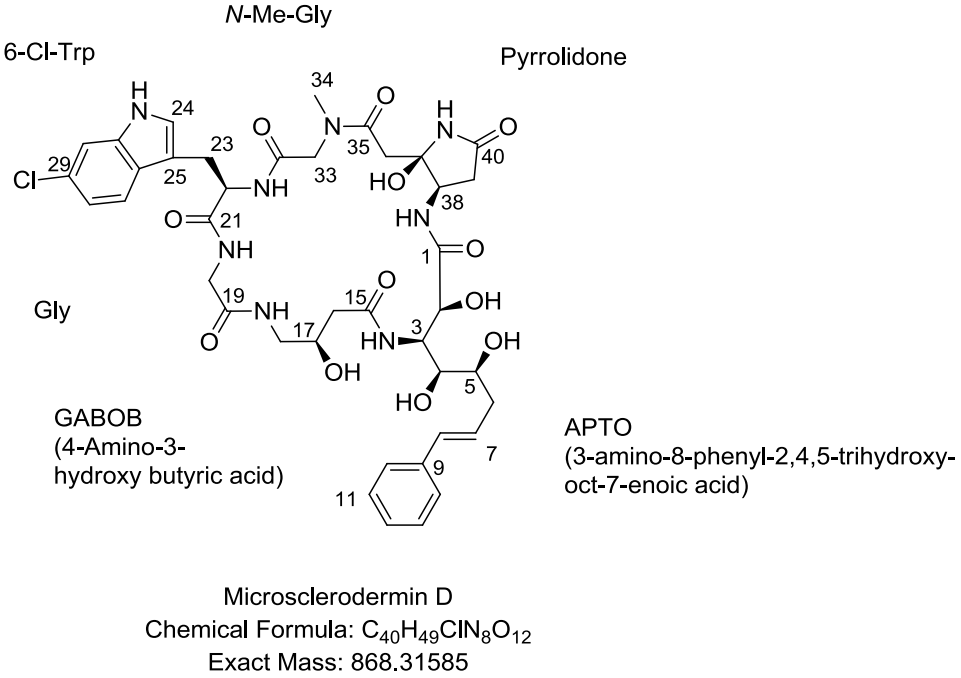


Figure S1: Structure and atom numbering of microsclerodermin D. Compound isolated from marine sponges of the genera Microscleroderma sp. and Theonella sp. and the myxobacterium Jahnella sp. MSr9139.

Table S1: NMR chemical shifts of microsclerodermin D.

Amino acid	Assignment	δ_C [a]	δ_H (mult., *J* in Hz) [b]	HMBC	ROESY
APTO	1	172.3			
	2	69.2	4.39 (d, 3.2)		
	3	53	4.13 (d, 11.3)		
	4	69.5	3.32 (m)		
	5	68.8	3.57 (m)		
	6	36.3	2.35 (m)		
	7	128.0	6.24 (m)		
	8	130.5	6.40 (d, 15.8)		
	9	137.2			
	10, 14	125.4	7.35 (d, 7.3)		
	11, 13	128.2	7.29 (t, 7.6)		
	12	126.5	7.19 (t, 7.3)		
	OH-2		6.58 (brs)		
	OH-4		4.17 (m)		
	OH-5		4.36 (d, 4.5)		
	NH-3		7.47 (d, 6.0)		
GABOB	15	172.3			
	16	40.7	2.43 (d, 14.0)		
			2.15 (d, 14.0)		

	17	66.8	3.72 (m)
	18	44.7	3.39 (m)
			2.61 (m)
	OH-17		4.88 (d, 4.8)
	NH-18		7.50 (m)
Gly	19	168.5	
	20	42.3	3.75 (d, 7.0)
			3.34 (d, 7.0)
	NH-20		8.54 (t, 6.0)
6-Cl-Trp	21	171.5	
	22	55.1	4.17 (m)
	23	25.8	3.10 (dd, 5.7, 14.7)
			2.98 (dd, 5.7, 14.7)
	24	124.6	7.26 (d, 1.8)
	25	109.5	
	26	125.6	
	27	119.4	7.52 (d, 8.3)
	28	118.4	7.00 (dd, 1.5, 8.3)
	29	c	
	30	110.7	7.37 (s)
	31	136.1	
	NH-22		8.64 (d, 4.3)
	NH-24		11.04 (brs)
N-Me-Gly	32	169.9	
	33	49.5	4.08 (d, 16.3)
			3.84 (d, 16.3)
	34	36.2	2.93 (s)
Pyrrolidone	35	170.2	
	36	38.6	2.84 (d, 17.1)
			2.69 (d, 17.1)
	37	85.2	
	38	50.3	4.47 (m)
	39	35.0	2.27 (m)
	40	172.3	
	NH-37		7.96 (brs)
	NH-38		7.56 (m)
	OH-37		c

[a] Recorded at 175 MHz, referenced to residual solvent DMSO-d_6 at 39.51 ppm. [b] Recorded at 700 MHz, referenced to residual solvent DMSO-d_6 at 2.50 ppm. [c] Not observed. [ov] overlapping signals.

NMR data of Dehydromicrosclerodermin D - acetonide

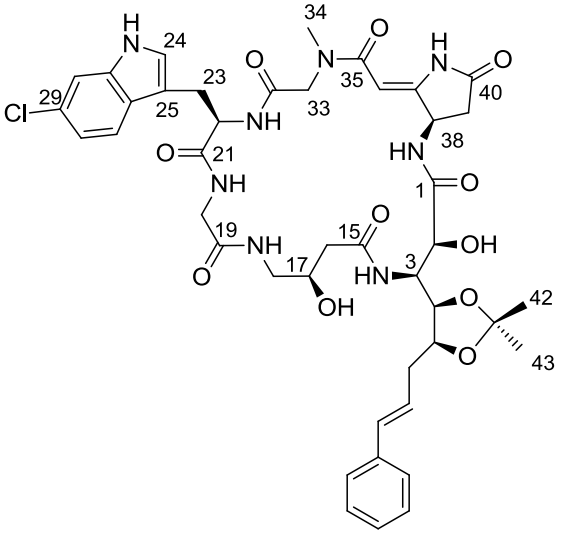


Figure S2: Structure and atom numbering of the acetonide of dehydromicrosclerodermin D. Derivative is based on microsclerodermin D as isolated from the myxobacterium Jahnella sp. MSr9139.

Table S2: NMR chemical shifts of dehydromicrosclerodermin D - acetonide.

Amino acid	Assignment	δ_C [a]	δ_H (mult., *J* in Hz) [b]
APTO	1	172.3	
	2	70.1	4.19 (d, 4.2)
	3	55.4	4.11 (d, 10.0)
	4	77.3	3.74 (ov)
	5	78.7	4.02 (m)
	6	37.1	2.32 (m)
			2.55 (m)
	7	126.5	6.24 (m)
	8	131.3	6.43 (d, 16.0)
	9	c	
	10, 14	125.4	7.38 (m)
	11, 13	128.2	7.30 (t, 7.6)
	12	126.5	7.20 (t, 7.2)
	OH-2		6.23 (ov)
	NH-3		7.09 (ov)
GABOB	15	c	
	16	41.1	2.22 (d, 14.8)
			2.01 (m)
	17	65.6	3.89 (m)
	18	44.7	3.23 (m)
			2.82 (m)
	OH-17		5.10 (d, 4.5)
	NH-18		7.35 (t, 6.5)
Gly	19	c	
	20	42.6	3.58 (m))

	NH-20		8.38 (t, 6.5)
6-Cl-Trp	21	c	
	22	55.2	4.11 (m)
	23	25.4	3.19 (ov)
			3.01 (dd, 5.1, 10.0)
	24	124.6	7.26 (d, 1.8)
	25	109.5	
	26	125.6	
	27	119.4	7.52 (d, 8.3)
	28	118.4	7.00 (dd, 1.5, 8.3)
	29	c	
	30	110.7	7.37 (s)
	31	136.1	
	NH-22		8.77 (d, 5.0)
	NH-24		11.04 (brs)
N-Me-Gly	32	c	
	33	50.5	3.44 (d, 15.7)
			4.53 (d, 15.7)
	34	36.8	3.06 (s)
Pyrrolidone	35	c	
	36	88.0	5.25 (s)
	37	c	
	38	45.4	5.21 (m)
	39	33.7	2.44 (m)
			2.74 (m)
	40	c	
	NH-37		c
	NH-38		8.36 (d, 8.8)

[a] Recorded at 175 MHz, referenced to residual solvent DMSO-d_6 at 39.51 ppm. [b] Recorded at 700 MHz, referenced to residual solvent DMSO-d_6 at 2.50 ppm. [c] Not observed. [ov] overlapping signals.

NMR data of Microsclerodermin L

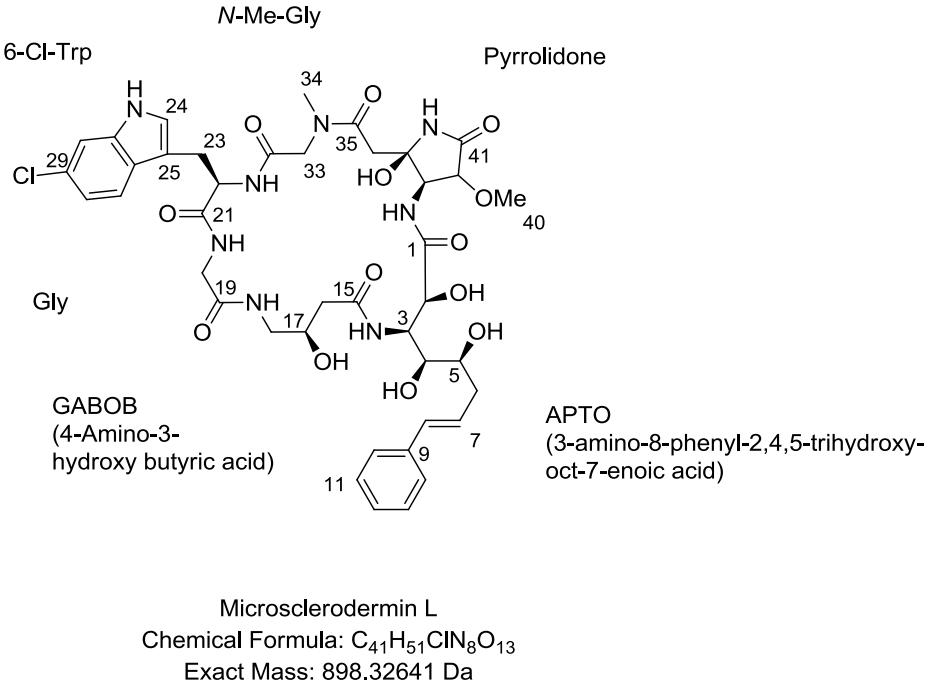


Figure S3: Structure and atom numbering of microsclerodermin L. Compound isolated from the myxobacterium Jahnella sp. MSr9139.

Table S3: NMR chemical shifts of microsclerodermin L.

Amino acid	Assignment	δ_C [a]	δ_H (mult., *J* in Hz) [b]	HMBC	ROESY
APTO	1	172.9			
	2	69.2	4.43 (brs)	1	
	3	53	4.19 (d, 11.3)	4	4
	4	69.7	3.32 (m)		
	5	68.7	3.59 (q, 6.5)		
	6	36.6	2.38 (m)	4, 5, 7, 8	
			2.33 (m)		
	7	128	6.26 (m)	6, 9	
	8	130.5	6.40 (d, 15.8)	6, 9, 10, 14	6
	9	137.2			
	10, 14	125.6	7.36 (d, 7.0)	8, 12	7, 8
	11, 13	128.3	7.29 (t, 7.6)	9, 10, 14	
	12	126	7.19 (t, 7.4)	10, 11, 13, 14	
	OH-2		6.57 (brs)		
	OH-4		4.85 (d, 4.9)		
	OH-5		4.30 (brs)		
	NH-3		7.42 (brs)		16
GABOB	15	172.1			
	16	40.6	2.43 (d, 14.0)	15	
			2.12 (d, 14.0)		
	17	66.6	3.77 (m)	16, 18	OH-17
	18	44.8	3.33 (m)	17	
			2.65 (m)		
	OH-17		4.85 (brs)		
	NH-18		7.43 (brs)		20

Gly	19	168.5			
	20	42.5	3.72 (d, 7.0) 3.36 (d, 7.0)	19, 21	
	NH-20		8.51 (t, 6.0)		22
6-Cl-Trp	21	171.6			
	22	55.2	4.17 (m)		NH-20
	23	25.8	3.10 (dd, 5.7, 15.0) 2.98 (dd, 5.7, 15.0)	21, 22, 24, 25, 26	
	24	124.7	7.26 (d, 2.0)	25, 31, 32	
	25	109.3			
	26	125.5			
	27	119.4	7.52 (d, 8.5)	31	
	28	118.5	7.00 (dd, 1.8, 8.5)	30, 26	
	29	c			
	30	110.7	7.37 (s)	27	
	31	136.3			
	NH-22		8.68 (d, 4.3)		22, 23, 24, 33
	NH-24		11.03 (d, 2.0)	24, 25, 31, 26	24, 30
N-Me-Gly	32	172.2			
	33	49.6	4.28 (d, 16.4) 3.69 (d, 16.4)	32, 34	
	34	35.9	2.92 (s)	33, 35	
Pyrrolidone	35	170.4			
	36	38.6	2.83 (d, 16.9) 2.61 (d, 16.9)	35, 37	38
	37	82.3			
	38	55.8	4.44 (d, 9.0)	39	NH-38
	39	79.0	4.04 (d, 8.8)	40, 41	NH-38
	40	56.6	3.39 (s)	39	
	41	171.0			
	NH-37		8.18 (brs)	37, 38, 39	
	NH-38		7.61 (d, 9.6)		2, 39
	OH-37		c		

[a] Recorded at 175 MHz, referenced to residual solvent DMSO-d_6 at 39.51 ppm. [b] Recorded at 700 MHz, referenced to residual solvent DMSO-d_6 at 2.50 ppm. [c] Not observed. [ov] overlapping signals.

NMR data of Microsclerodermin M

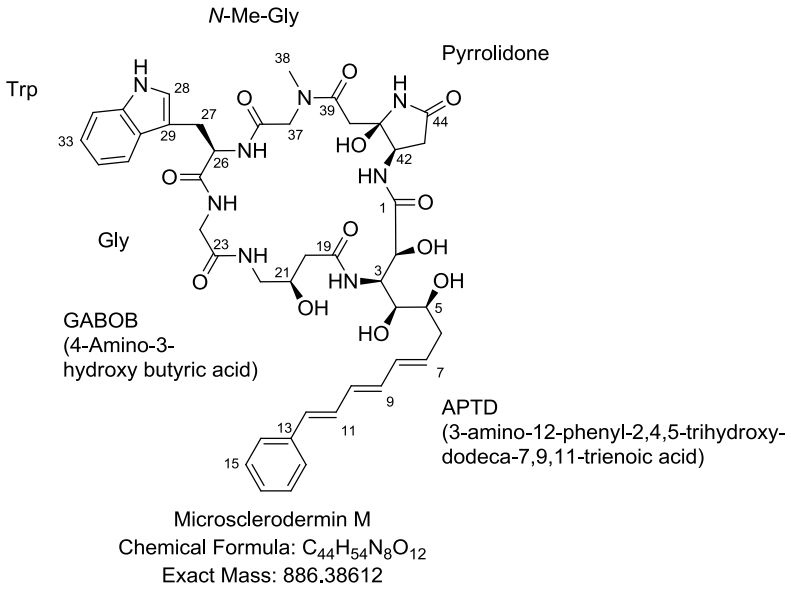


Figure S4: Structure and atom numbering of microsclerodermin M. Compound isolated from the myxobacterium Sorangium cellulosum So ce38.

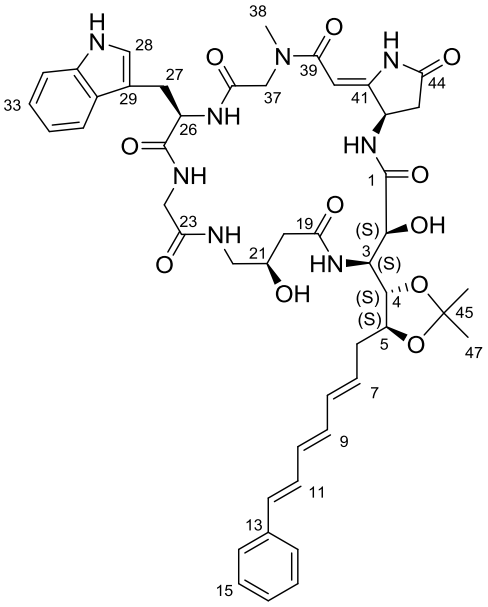
Table S4: NMR chemical shifts of microsclerodermin M.

Amino acid	Assignment	δ_C [a]	δ_H (mult., *J* in Hz) [b]	HMBC	ROESY
APTD	1	172.5			
	2	69.3	4.38 (d, 5.4)	3	3, OH-2
	3	53.0	4.11 (ov)	4, 5, 19	OH-5, NH-3
	4	69.8	3.27 (ov)	2, 3	3, 7, OH-2, OH-3
	5	68.8	3.52 (ov)	4, 6	7, 8
	6	36.4	2.28 (ov)	4, 5, 7, 8	4, 5, 7, 8
	7	133.0	5.74 (dt, 15.0, 7.4)	6, 9	4, 5, 6, 9
	8	131.7	6.17 (dd, 10.0, 15.1)	9, 10	6, 11
	9	133.7	6.34 (ov)	7, 8, 10, 11	7, 11
	10	130.7	6.33 (ov)	8, 11, 12	12
	11	129.2	6.92 (dd, 9.6, 15.6)	9, 10, 13	9, 18
	12	131.3	6.56 (d, 15.6)	10, 14, 18, 13	10, 14
	13	137.0			
	14, 18	125.7	7.44 (d, 7.6)	12, 16	11, 12
	15, 17	128.4	7.31 (d, 8.3)	14, 18	14
	16	127.1	7.21 (d, 7.1)	14, 15, 17, 18	14, 18
	OH-2		6.25 (d, 5.6)	2, 3	2, 4, NH-42
	OH-4		4.34 (d, 4.7)	5, 6	3, 5
	OH-5		4.52 (d, 9.0)	3, 4, 5	3, 4
	NH-3		7.43 (d, 12.6)	2, 3	4, 5, 20, OH-21
GABOB	19	172.6			
	20	40.7	2.44 (d, 14.0) 2.15 (ov)	19, 21, 22	NH-3

	21	66.9	3.73 (ov)	21, 23	20, 22, OH-21, NH-22
	22	44.8	3.37 (ov) 2.64 (ov)		21, OH-21
	NH-22		7.49 (d, 6.7)		20, 21, 22
	OH-21		4.86 (d, 5.0)	20, 21, 22	20, 21, 22, NH-3
Gly	23	168.7			
	24	42.4	3.76 (d, 7.6) 3.36 (d, 5.7)	23, 25	NH-24
	NH-24		8.56 (t, 5.7)		24, 26, NH-22
Trp	25	171.7			
	26	55.2	4.19 (m)	25, 27, 29, 36	27, 30, NH-24
	27	26.2	3.11 (dd, 5.6, 14.6) 2.98 (dd, 5.6, 14.6)	25, 26, 29	26, 28, 30, NH-26
	28	123.5	7.21 (d, 6.6)	29, 34	26, 27, NH-26, NH-28
	29	109.6			
	30	126.5			
	31	118.0	7.52 (d, 7.7)	29, 32	26, 27, 32
	32	118.1	6.98 (t, 7.3)	33, 34	31, 34
	33	120.8	7.06 (d, 7.3)	30, 31, 35	31, 34
	34	111.1	7.32 (d, 7.9)	30, 32, 35	32
	35	135.9			
	NH-28		10.87 (brs)	28, 29, 35, 30	28, 34
	NH-26		8.65 (d, 4.2)	26, 27	26, 27, 28, 37
***N*-Me-Gly**	36	169.9			
	37	49.6	4.10 (d, 15.4) 3.82 (d, 15.4)	39, 36	38, NH-26
	38	36.2	2.94 (s)	39	37
Pyrrolidone	39	170.4			
	40	38.7	2.85 (d, 17.0) 2.70 (d, 17.0)	39, 41, 42	42, OH-41, NH-41
	41	85.5			
	42	50.4	4.47 (ov)	40, 42, 43	40
	43	34.9	2.28 (ov)	41, 44	42, OH-41, NH-42
	44	172.9			
	OH-41		6.07 (s)	40, 42	40, 43, NH-41, NH-42
	NH-41		7.97 (s)	41, 42, 43	40, 42, OH-41
	NH-42		7.54 (ov)	42	42, 43, OH-41

[a] Recorded at 175 MHz, referenced to residual solvent DMSO-d_6 at 39.51 ppm. [b] Recorded at 700 MHz, referenced to residual solvent DMSO-d_6 at 2.50 ppm. [c] Not observed. [ov] overlapping signals.

NMR data of Dehydromicrosclerodermin M - acetonide



Dehydromicrosclerodermin M
Acetonide
Chemical Formula: $C_{47}H_{56}N_8O_{11}$
Exact Mass: 908.40685

Figure S5: Structure and atom numbering of the acetonide of dehydromicrosclerodermin M. Derivative is based on microsclerodermin M as isolated from the myxobacterium Sorangium cellulosum So ce38.

Table S5: NMR chemical shifts of microsclerodermin M - acetonide.

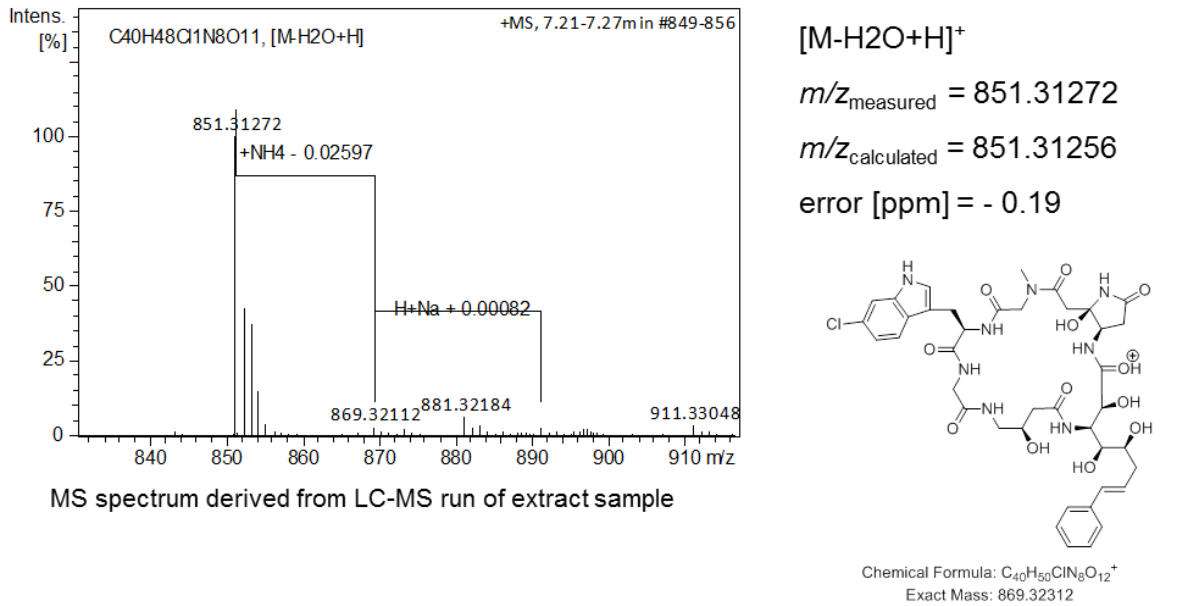
Amino acid	Assignment	δ_C [a]	δ_H (mult., *J* in Hz) [b]
APTD	1	172.5	
	2	69.9	4.17 (ov)
	3	53.1	4.28 (m)
	4	77.0	3.69 (dd, 7.0, 9.9)
	5	78.6	3.96 (m)
	6	36.7	2.47, 2.24 (ov)
	7	133.0	5.75 (m)
	8	131.7	6.17 (dd, 9.7, 15.0)
	9	133.7	6.34 (ov)
	10	130.7	6.33 (ov)
	11	129.2	6.92 (dd, 9.6, 15.6)
	12	131.3	6.56 (d, 15.6)
	13	137.0	
	14, 18	125.7	7.44 (d, 7.6)
	15, 17	128.4	7.31 (ov)
	16	127.1	7.21 (t, 7.3)
	45	*c*	
	46	26.6	1.35 (s)
	47	26.6	1.35 (s)
	OH-2		6.25
	NH-3		7.43
GABOB	19	172.6	
	20	40.3	2.22 (ov)

			2.00 (dd, 7.7, 10.8)
	21	65.2	3.87 (m)
	22	44.8	3.23 (ov)
			2.82 (m)
	NH-22		7.37
	OH-21		4.86
Gly	23	168.7	
	24	41.9	3.59 (m)
	NH-24		8.40
Trp	25	171.7	
	26	54.8	4.14 (m)
	27	25.1	3.21 (dd, 3.3, 10.3)
			3.00 (dd, 3.3, 10.3)
	28	123.5	7.21 (d, 5.3)
	29	109.6	
	30	126.5	
	31	118.0	7.52 (d, 7.7)
	32	118.1	6.98 (t, 7.3)
	33	120.8	7.06 (t, 7.3)
	34	111.1	7.32 (d, 7.9)
	35	135.9	
	NH-28		10.87 (brs)
	NH-26		8.65
***N*-Me-Gly**	36	169.9	
	37	49.7	4.54 (d, 11.5)
			3.43 (ov)
	38	36.4	3.05 (s)
Pyrrolidone	39	170.4	
	40	87.3	5.23 (ov)
	41	*c*	
	42	50.4	3.69 (m)
	43	33.1	2.45, 2.75 (ov)
	44	172.9	
	NH-41		8.35
	NH-42		*c*

[a] Recorded at 175 MHz, referenced to residual solvent DMSO-d_6 at 39.51 ppm. [b] Recorded at 700 MHz, referenced to residual solvent DMSO-d_6 at 2.50 ppm. [c] Not observed. [ov] overlapping signals.

Analytical data for microsclerodermins and pedeins

Microsclerodermin D



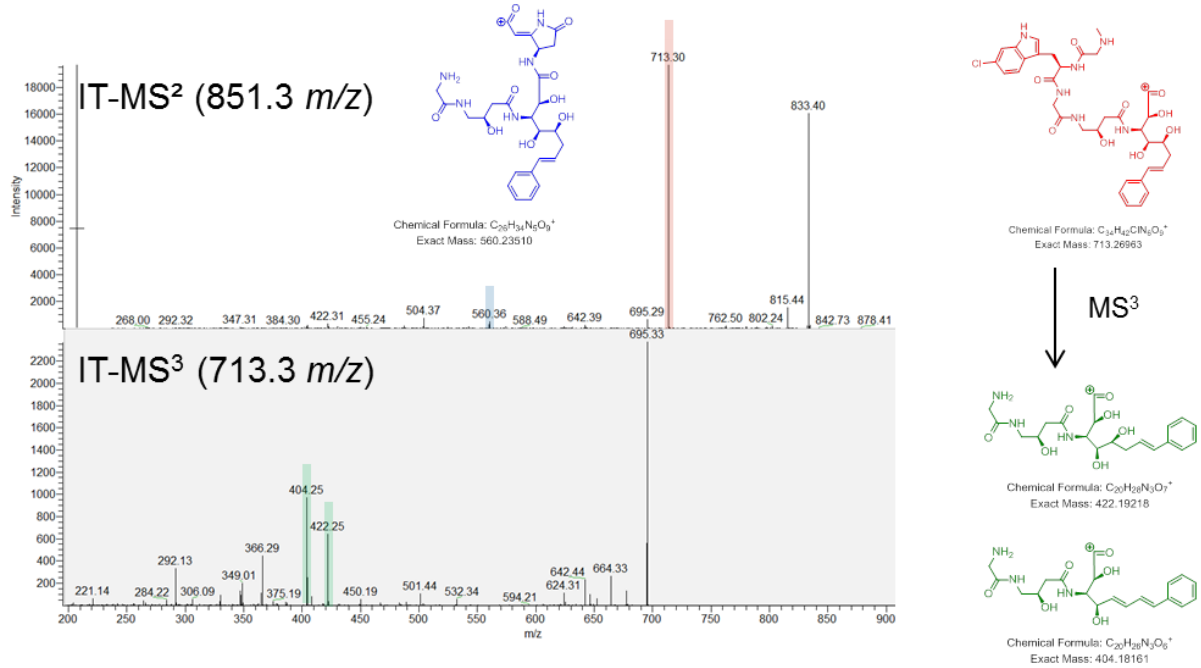
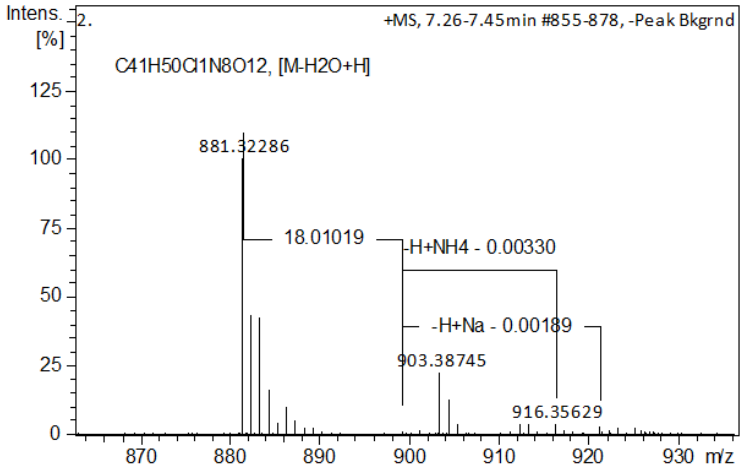


Figure S6: Accurate m/z measurement and CID fragmentation pattern of microsclerodermin D.

Microsclerodermin L



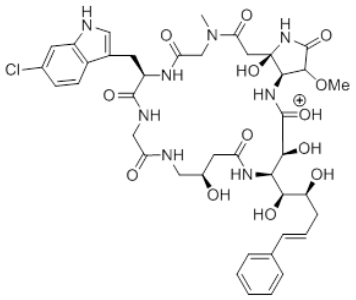
MS spectrum derived from LC-MS run of extract sample

$[M\text{-}H2O+H]^+$

$m/z_{measured}$ = 881.32286

$m/z_{calculated}$ = 881.32312

error [ppm] = + 0.30



Chemical Formula: $C_{41}H_{52}ClN_8O_{13}^+$
Exact Mass: 899.33369

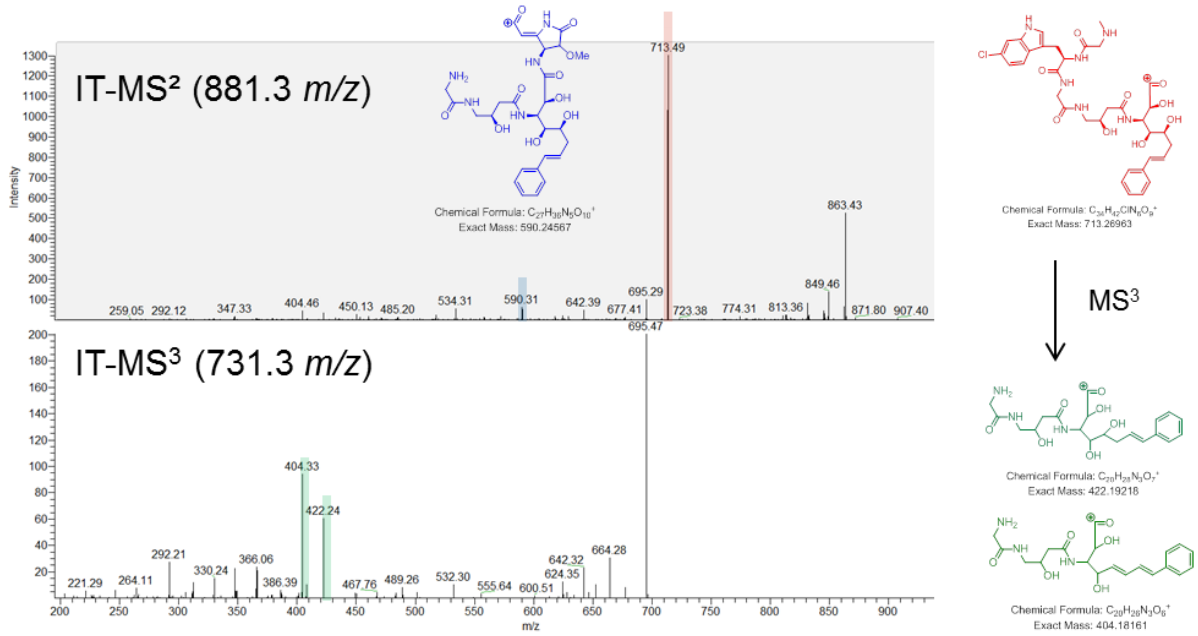
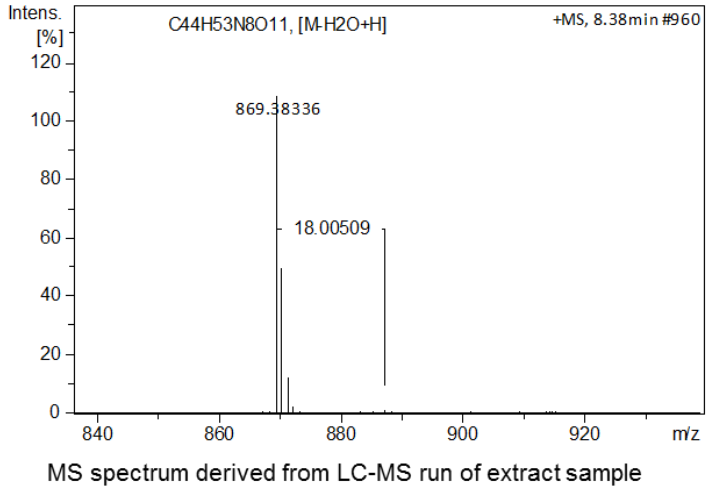


Figure S7: Accurate m/z measurement and CID fragmentation pattern of microsclerodermin L.

Microsclerodermin M

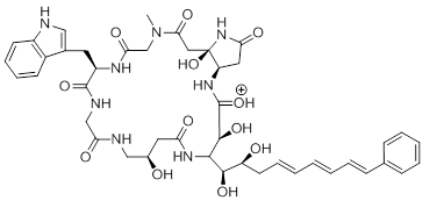


$[M\text{-}H_2O+H]^+$

$m/z_{measured}$ = 869.38336

$m/z_{calculated}$ = 869.38283

error [ppm] = - 0.61



Chemical Formula: $C_{44}H_{55}N_8O_{12}^+$
Exact Mass: 887.39340

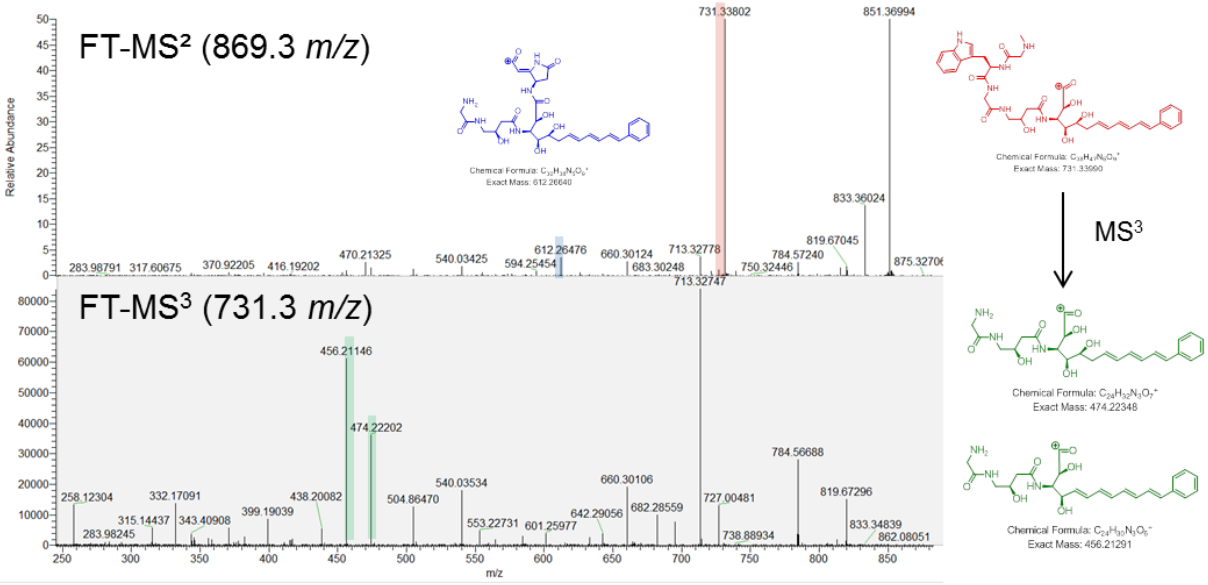
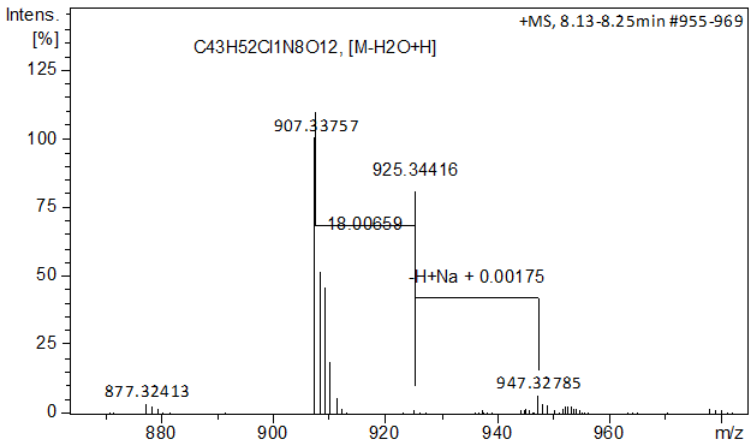


Figure S8: Accurate m/z measurement and CID fragmentation pattern of microsclerodermin M.

Pedein A



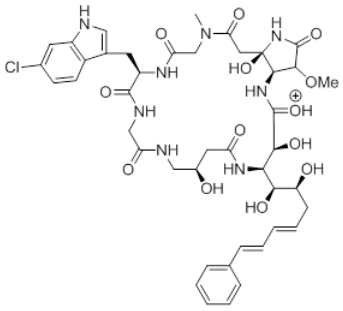
MS spectrum derived from LC-MS run of extract sample

$[M\text{-}H2O+H]^+$

$m/z_{measured}$ = 907.33757

$m/z_{calculated}$ = 907.33877

error [ppm] = - 1.32



Chemical Formula: $C_{43}H_{54}ClN_8O_{13}^+$
Exact Mass: 925.34934

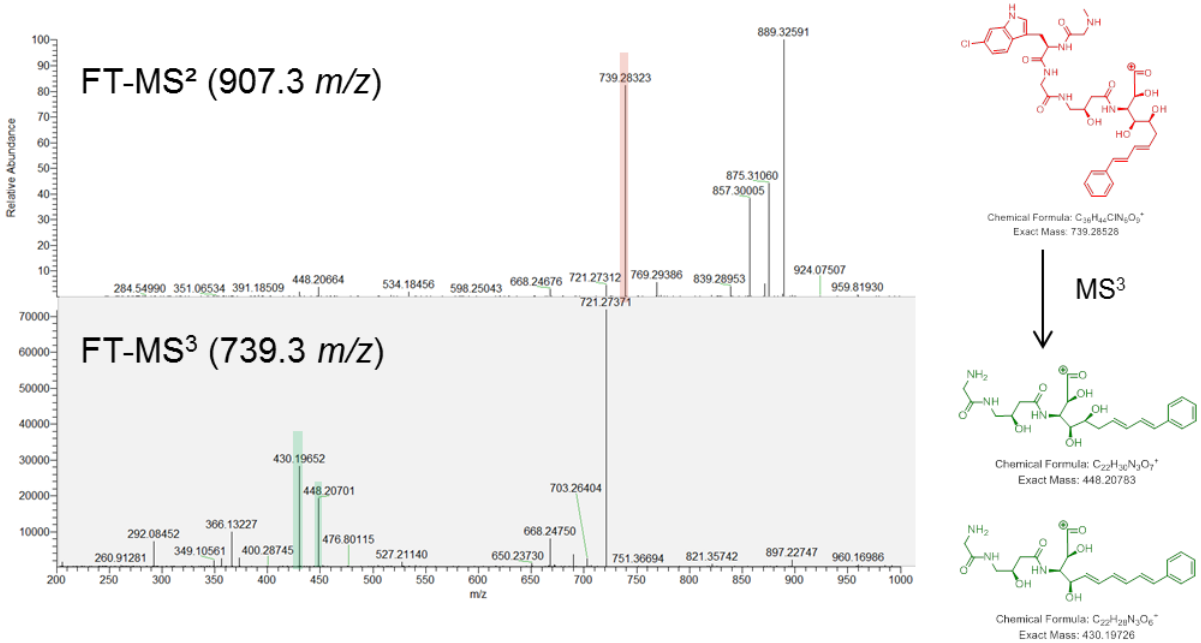
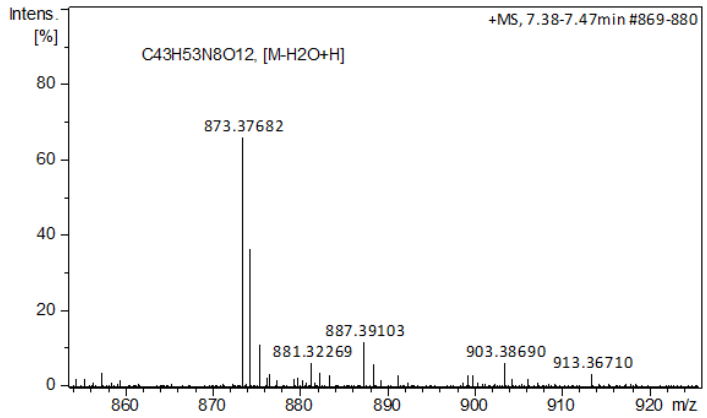


Figure S9: Accurate m/z measurement and CID fragmentation pattern of pedein A.

Pedein B



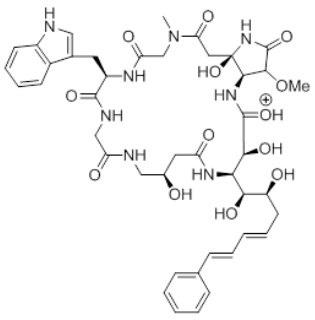
MS spectrum derived from LC-MS run of extract sample, thus coeluting with other derivatives of microsclerodermin.

$[M\text{-}H_2O+H]^+$

$m/z_{measured}$ = 873.37682

$m/z_{calculated}$ = 873.37775

error [ppm] = 1.07



Chemical Formula: $C_{43}H_{55}N_8O_{13}^+$
Exact Mass: 891.38831

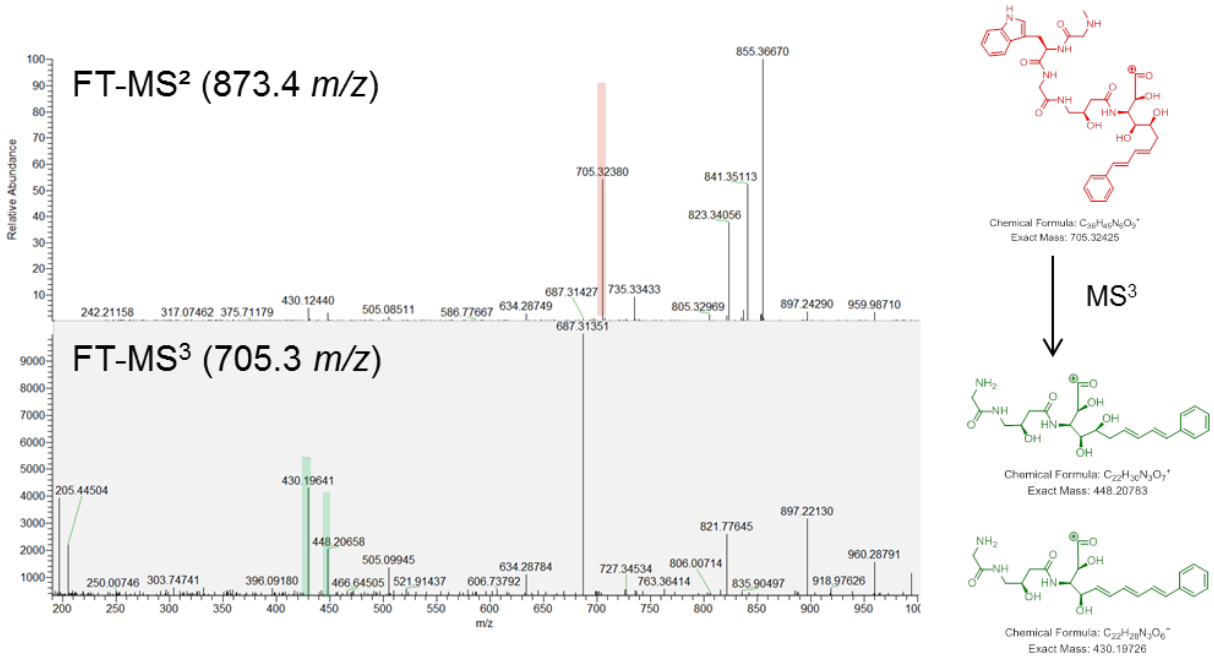


Figure S10: Accurate m/z measurement and CID fragmentation pattern of pedein B.

Analysis of the *msc* gene clusters in So ce38 and MSr9139

The regions of interest as well as 20 kb flanking regions were searched for open reading frames (ORFs) using Glimmer 3.0 and subsequently subjected to an automatic annotation by the antiSMASH software tool[1,2]. The *msc*-locus in So ce38 (Table S6) spans 58,049 bp and covers nine modules within the genes *mscA* to *mscI* with a GC content of 74.7 % (EMBL accession no. XXXX). The 5'-end of the gene cluster covers a major facilitator superfamily transporter (MFS-type, *mscK*) and a type II-thioesterase (*mscJ*) located approximately 2.5 kb upstream to *msc*A.

In the case of *Jahnella* sp. MSr9139 the cluster has a size of 62,268 bp and is thereby approximately 4.2 kb larger. We find the same genes as in So ce38 starting from *mscA* to *mscI* (Table S6, EMBL accession no. XXXX). There are three additional auxiliary genes named *mscL*, *mscM* and *mscN*. The first one encodes a halogenase specific for tryptophan found at the end of the cluster. This is consistent with the observation that a chlorinated tryptophan residue is found for the microsclerodermins isolated from *Jahnella* specimen. The genes *mscM* and *mscN* are located within the cluster (between *mscD* and *mscE*). They are encoding two proteins with significant similarity to an α-ketoglutarate dependent dioxygenase and a methyltransferase, respectively. The pairwise similarity of both clusters is 74.2 % on DNA-level. Both *msc* clusters have a comparable GC content of around 72 %.

Table S6: Genes involved in microsclerodermin biosynthesis as annotated in the strains Sorangium cellulosum *So ce38 and* Jahnella *sp. MSr9139*

Sorangium cellulosum So ce38					*Jahnella* sp. MSr9139				
Gene	Start	End	Length [bp]	Function[1]	Start	End	Length [bp]	Function[1]	pairwise identity [%]
mscA	3859	13686	9828	CoA-Lig-KR*-ACP-KS-AT-DH-KR-ACP	4124	14731	10608	CoA-Lig-KR* -MT-ACP-KS-AT-DH-KR-ACP	74.7
mscB	13709	16321	2613	KS-AT*	14718	17354	2637	KS-AT*	79.7
mscC	16327	20982	4656	KS-AT-DH-KR-ACP	17341	21990	4650	KS-AT-DH-KR-ACP	81.5
mscD	20995	23697	2703	KS-AT*	22000	24546	2547	KS-AT*	78.7
mscE	23727	25067	1341	Putative amidohydrolase	27907	26747	1161	Putative amidohydrolase	83.9
mscF	25335	32156	6822	PCP-ATT-MOX-C-A-PCP	28287	34856	6570	PCP-AAT-MOX-C-A-PCP	82.4
mscG	32158	36804	4647	KS-AT-KR-ACP	34858	39393	4536	KS-AT-KR-ACP	80.9
mscH	36844	49269	12426	C-A-M-PCP-C-A-PCP-E-C-A-PCP	39472	51792	12321	C-A-M-PCP-C-A-PCP-E-C-A-PCP	84.5
mscI	49344	58049	8706	C-A-PCP-KS-AT-KR-ACP-TE	51789	60668	8880	C-A-PCP-KS-At-KR-ACP-TE	83.9
mscJ	1473	2246	774	Thioesterase typeII	3006	3794	789	Thioesterase typeII	63
mscK	1248	1	1248	MFS-transporter	1350	1	1350	MFS-transporter	55.9
mscL	-	-	-	-	62268	60661	1608	Trp-halogenase	-
mscM	-	-	-	-	24583	25899	1317	Fe(II)/α-ketoglutarate dependent oxygenase	-

mscN	-	-	-	-	25881	26714	834	Methyltransferase	-

[1] A: adenylation domain, AAT: aminotransferase, ACP: acyl-carrier-protein domain AT: acyltransferase domain, C: condensation domain, CoA-Lig: CoA-Ligase, DH: dehydratation domain, E: epimerase, KR: ketoreductase domain, KS: ketosynthase domain, MT: methyltransferase domain, MOX: monooxygenase, PCP: peptidyl-carrier-protein domain, TE: thioesterase domain, *: inactive domain

Table S7: Proteins involved in microsclerodermin biosynthesis in the strains Sorangium cellulosum *So ce38 and* Jahnella *sp. MSr9139.*

Sorangium cellulosum So ce38			*Jahnella* sp. MSr9139		
Protein	Length [aa]	Domains and position in sequence[a]	Length [aa]	Domains and position in sequence[a]	Identity [%]
MscA	3275	CoA-Lig (264-701), KR*(1034-1197), ACP (1348-1411), KS (1441-1828), AT (1976-2286), DH (2342-2505), KR' (2870-3047), ACP' (3149-3214)	3535	CoA-Lig (215-649), KR*(1032-1126), MT (1229-1509), ACP (1613-1676), KS (1712-2147), AT (2244-2555), DH (2610-2774), KR*' (3132-3309), ACP' (3411-3476)	65.4
MscB	870	KS (27-451), AT* (548-772)	878	KS (39-464), AT* (561-792)	73.4
MscC	1551	KS (36-460), AT (557-859), DH* (956-1076), KR (1158-1336), ACP (1436-1505)	1549	KS (39-464), AT (561-865), DH* (956-1076), KR (1157-1335), ACP (1436-1499)	75.0
MscD	900	KS (36-461), AT* (565-678)	848	KS (36-461), AT* (564-675)	72.0
MscE	446	Putative amidohydrolase	386	Putative amidohydrolase	82.9
MscF	2273	PCP (27-98), AMT (329-660), MOX (828-1127), C (1185-1529), A (1673-2082), PCP' (2169-2237)	2189	PCP (4-75), AMT (280-614), MOX (758-1057), C (1105-1397), A (1594-2000), PCP' (2088-2149)	76.5
MscG	1548	KS (14-439), AT (534-828), KR (1156-1330), ACP (1441-1509)	1511	KS (14-439), AT (531-850), KR (1136-1312), ACP (1404-1472)	72.1
MscH	4141	C (76-377), A (564-966), MT (1037-1256), PCP (1469-1531), C' (1554-1850), A' (2037-2426), PCP' (2515-2574), E (2591-2905), C'' (3074-3374), A'' (3558-3967), PCP'' (4054-4121)	4106	C (48-346), A (534-936), MT (1007-1225), PCP (1442-1505), C' (1526-1827), A' (2013-2405), PCP' (2492-2551), E (2568-2872), C'' (3043-3343), A'' (3527-3932), PCP'' (4019-4083)	78.8
MscI	2904	C (48-346), A (533-936), PCP (1023-1087), KS (1111-1535), AT (1638-1936), KR (2266-2465), ACP (2549-2612), TE (2634-2888)	2945	C (77-375), A (563-965), PCP (1053-1116), KS (1150-1573), AT (1676-1970), KR (2309-2509), ACP (2597-2660), TE (2683-2945)	77.8
MscJ	257	Thioesterase type II	263	Thioesterase type II	43.5
MscK	415	Major Facilitator Superfamily (MFS) transporter	450	Major Facilitator Superfamily (MFS) transporter	24.5
MscL	-	-	535	Trytophan halogenase	-
MscM	-	-	438	Fe(II)/α-ketoglutarate dependent oxygenase	-
MscN	-	-	277	Methyltransferase	-

[a] A: adenylation domain, AMT: aminotransferase, ACP: acyl-carrier-protein domain AT: acyltransferase domain, C: condensation domain, CoA-Lig: coenzyme A Ligase, DH: dehydratation domain, E: epimerase, KR: ketoreductase domain, KS: ketosynthase domain, MT: methyltransferase domain, MOX: monooxygenase, PCP: peptidyl-carrier-protein domain, TE: thioesterase domain, *: inactive domain

AT domains

Substrate specificity of acyl transferase (AT) domains was determined using conserved motif analysis by aligning the AT domains from both clusters to a reference AT from *E.coli* FAS, 1MLA (PDB 1MLA, UniProtKB P0AAI9) (Table S8).[3] This data shows that most of the listed AT-

domains contain the GxSxG consensus motif with the catalytic residue in the center, except MscB in which the catalytic serine residue is replaced by a glycine. The MscD-AT domains are considered inactive as they are heavily truncated and do not align to the canonical motifs of AT domains. Active site analysis was used to predict which extender unit is recruited by a distinct AT domain. The main discriminant between using a methylmalonyl-CoA (mm-CoA) or malonyl-CoA (m-CoA) is the amino acid position 200 in 1MLA, where Ser200 supports mm-CoA and Phe200 supports m-CoA. The *in silico* prediction for microsclerodermin production is in all cases m-CoA and thereby in agreement with the chemical structure observed. We were not able to predict the incorporation of a hydroxymalonyl-CoA substrate for MscC-AT.

Table S8: Active site analysis of the acyl transferase domains (AT) from S. cellulosum *So ce38 and* Jahnella *sp. MSr9139 according to Yadav* et al.[3]

Domain	observed	predicted	11	63	90	91	92	93	94	117	200	201	231	250	255	15	58	59	60	61	62	70	72	197	198	199
1MLA	M	M	Q	Q	G	H	S	L	G	R	S	H	N	Q	V	T	K	T	W	Q	T	S	A	S	V	P
MscA-AT$_{So\ ce38}$	M	M	Q	Q	G	H	S	V	G	R	F	H	N	H	V	Q	R	T	E	Y	T	E	A	S	H	A
MscA-AT$_{Jahnella\ sp.}$	M	M	Q	Q	G	H	S	V	G	R	F	H	N	H	V	Q	R	T	E	Y	T	E	A	S	H	A
MscB-AT$_{So\ ce38}$	inactive	-	E	Q	G	Q	G	A	G	R	-	-	S	D	D	T	Q	T	A	F	T	Q	A	-	-	-
MscB-AT$_{Jahnella\ sp.}$	inactive	-	Q	Q	G	L	G	V	G	R	-	-	S	Q	D	L	Q	A	A	F	A	E	A	-	-	-
MscC-AT$_{So\ ce38}$	Hydroxy malonate	M	Q	Q	G	H	S	I	G	R	F	H	N	H	V	Q	Q	T	A	F	T	E	A	S	H	A
MscC-AT$_{Jahnella\ sp.}$	Hydroxy malonate	M	Q	Q	G	H	S	I	G	R	F	H	N	H	V	F	Q	T	A	F	A	E	A	S	H	A
MscG-AT$_{So\ ce38}$	M	M	Q	Q	G	H	S	I	G	R	F	H	N	H	V	H	D	T	A	I	A	E	A	S	H	A
MscG-AT$_{Jahnella\ sp}$	M	M	Q	Q	G	H	S	V	G	R	F	H	N	H	V	H	D	T	A	L	A	G	A	S	H	A
MscI-AT$_{So\ ce38}$	M	M	Q	Q	G	H	S	V	G	R	F	H	N	H	V	Y	Q	T	R	L	T	E	A	S	H	A
MscI-AT$_{Jahnella\ sp}$	M	M	Q	Q	G	H	S	V	G	R	F	H	N	H	V	H	Q	T	R	L	A	E	A	S	H	A

KS domains

The ketosynthase (KS) domains of both msc biosynthetic gene clusters were extracted and analyzed using the NaPDoS web tool.[4] The result is depicted as a phylogenetic tree in Figure S11 and allows easy annotation of the KS subtypes found in the msc cluster.

The postulated iterative function of MscB is supported by grouping of these domains at the interface between modular (ochre) and iterative (green) domains. The remaining, functional domains MscA, MscG, and MscI are correctly grouped in the PKS/NRPS hybrid section. MscA is characterized as a hybrid domain because of its linkage to the first module with its uncommon starter unit.

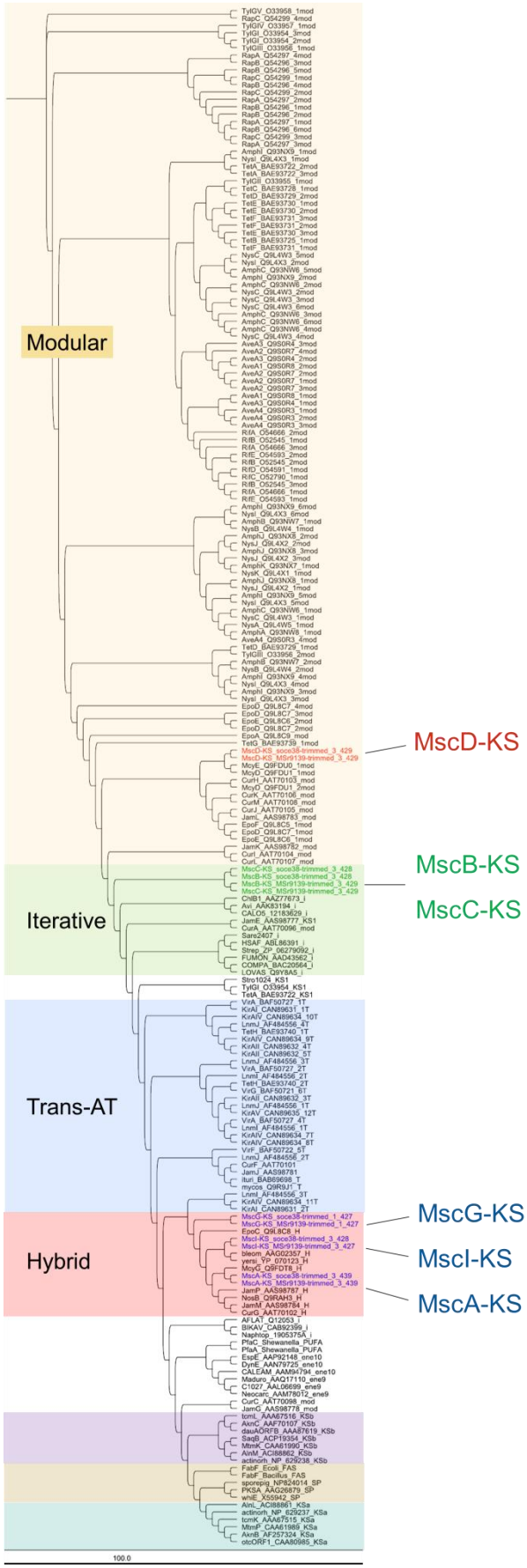


Figure S11: Phylogenetic tree highlighting different KS domain types and the KS domains found in the msc biosynthetic gene cluster in So ce38 and MSr9139.

DH domains

Both *msc* clusters contain two dehydration domains present in the proteins MscA and MscC. Based on the proposed biosynthesis of microsclerodermin the DH-domain in MscA must be active. However, it does not exhibit the common consensus-motif $H(X_3)G(X_4)P$ of DH-domains as known from other modular PKS-systems (Table S9). Here, the central glycine is replaced by a valine. This substitution is obviously not affecting the enzyme function. The DH-domains from MscC are considered to be inactive as they do not contain a conserved motif at all. The domains are truncated with a protein length of around 100 amino acids.

Table S9: Active site analysis of the dehydratase domains (DH) of the msc *biosynthetic gene cluster.*

Consensus	H	x	x	x	G	x	x	x	x	P
MscA-DH$_{So\ ce38}$	H	A	E	G	V	L	V	A	R	P
MscA-DH$_{\textit{Jahnella}\ sp}$	H	V	E	G	V	V	V	A	R	P

KR domains

The *msc* cluster from both myxobacteria encodes five ketoreductase domains. The KR domains may be classified as A- and B-type by analyzing the amino acid sequence of two regions within the amino acid sequence.[5] However, it was shown that a reliable prediction of the stereogenic center that emerges upon KR reduction of a keto-function is governed by additional, unknown interactions.[6] Nevertheless, we attempted to classify the domains in this work. The occurrence of a LDD motif is a good indicator for a B-type KR domain. This is sometimes supported by the presence of a Pro144 and an Asp148 according to the numbering based on the DEBS2 KR domain (GenBank #X62569.1). In case of A-type KR domains a highly conserved Trp141 is present. Moreover, a correct co-factor binding motif is mandatory for the enzyme function. Based on this, MscA-KR1 of both strains is predicted to be inactive as it does not exhibit the co-factor binding motif GxGxxGxxxA (Table S10). Additionally, the MscA-KR1 domain of MSr9139 is heavily truncated. The remaining KR domains feature the correct binding motif.

Table S10: Analysis of the co-factor binding motif of the ketoreductase domains (KR) within the msc *biosynthetic gene cluster.*

Domain	G	x	G	x	x	G	x	x	x	A	motif
MscA-KR1 $_{So\ ce38}$	D	L	G	G	V	S	L	Q	L	L	inactive
MscA-KR1 $_{MSr9139}$	A	L	G	S	Y	E	-	-	-	-	inactive
MscA-KR2 $_{So\ ce38}$	G	L	G	A	L	G	R	R	V	A	ok
MscA-KR2 $_{MSr9139}$	G	L	G	A	L	G	R	R	V	A	ok
MscC-KR$_{So\ ce38}$	G	T	G	A	L	G	L	H	I	A	ok
MscC-KR$_{MSr9139}$	G	I	G	A	L	G	L	H	T	A	ok
MscG-KR$_{So\ ce38}$	G	L	G	G	A	G	L	G	I	A	ok
MscG-KR$_{MSr9139}$	G	L	G	G	V	G	L	A	I	A	ok
MscI-KR$_{So\ ce38}$	G	L	G	R	I	G	L	C	L	A	ok
MscI-KR$_{MSr9139}$	G	L	G	R	I	G	L	C	L	A	ok

The two amino acid regions for KR domain classification are shown in Table S11. MscA-KR2 and MscC-KR1 can be classified as B-type and A-type, respectively. The LDD-motif together with an Asp148 is found in MscA-KR2 whereas the highly conserved Trp141 specifies MscC-KR. One stereogenic center originating from the A-type MscC-KR is still present in the final product and should exhibit *S*-configuration. This is in agreement with the configuration of the last hydroxyl group reported for the microsclerodermin side chain. The KR domains of MscG do only feature a LDD-like motif but no additional identifier in the second region. It remains unclear whether this domain is active at all. MscI-KR produces an *R*-configured hydroxyl moiety as identified by advanced Marfey analysis of microsclerodermin (GABA subunit). However, there is no common consensus that allows a classification. Moreover, the KR domains exceed the common KR domain length of approximately 180 amino acids by additional 30 amino acids.

Table S11: Analysis of the ketoreductase domains (KR) within the msc *biosynthetic gene cluster*

	region 88-103																region 134-149																KR type
MscA-KR2 $_{So\ ce38}$	H	L	A	G	A	L	D	D	G	V	L	L	Q	Q	S	W	F	S	S	I	A	S	V	L	G	S	A	A	Q	G	N	Y	B-type
MscA-KR2 $_{MSr9139}$	H	L	A	G	S	L	D	D	G	V	L	V	Q	Q	S	W	F	S	S	I	A	S	I	L	G	S	A	G	Q	G	N	Y	B-type
MscC-KR$_{So\ ce38}$	H	A	A	G	V	A	G	A	R	D	L	S	A	L	D	A	T	S	S	I	A	S	L	W	G	S	R	G	Q	A	H	Y	A-type
MscC-KR$_{MSr9139}$	H	A	A	G	V	G	G	Y	C	E	L	S	R	L	D	A	Y	S	S	I	A	S	L	W	G	S	R	G	Q	A	H	Y	A-type
MscG-KR$_{So\ ce38}$	Y	A	A	G	V	D	D	P	A	L	I	G	-	G	I	G	S	S	S	L	S	A	I	L	G	G	R	G	L	G	A	Y	B-type
MscG-KR$_{MSr9139}$	Y	A	A	A	I	D	D	A	G	L	L	G	-	A	V	G	G	S	S	L	S	A	I	L	G	G	R	G	L	A	A	Y	B-type
MscI-KR$_{So\ ce38}$	H	A	A	G	A	R	G	D	G	T	F	M	V	P	L	A	L	S	S	T	S	A	I	L	G	G	L	G	L	G	P	Y	unknown
MscI-KR$_{MSr9139}$	H	A	A	G	A	R	G	D	G	T	F	M	T	S	L	A	L	S	S	T	S	A	I	L	G	G	L	G	L	G	P	Y	unknown

A domains

Three genes of microsclerodermin biosynthesis (*mscF*, *mscH* and *mscI*) harbor at least one NRPS module. In order to determine the A domain specificity of each module the respective amino acid sequences were analyzed using the NRPSpredictor2 tool.[7] The results are listed in Table S12. The predicted specificity of each A-domain is consistent with the observed amino acid incorporated into the microsclerodermin scaffold.

Table S12: Prediction of A-domain specificity for each A-domain present in the msc *biosynthetic gene cluster using the NRPSpredictor2 software-tool.*[7]

domain	amino-acid specificity	
	So ce38	MSr9139
MscF-A	asn	asn
MscH-A1	gly	gly
MscH-A2	trp	trp
MscH-A3	gly	gly
MscI-A	gly	gly

C domains

All C domains of the *msc* biosynthetic gene cluster were analyzed using the NaPDoS web tool.[4] Especially the hybrid module MscF is of interest as the incorporated *S*-asparagine is (a) cyclized to form and pyrrolidone moiety and (b) converted to the *R*-asparagine. The mechanism behind these reactions is not understood yet. Analysis of the respective C domain, MscF-C1, characterizes the domain as a common PKS/NRPS hybrid domain (Figure S12). Hence, the observed cyclization as well as inversion of asparagine is not done by the C domain. It is more likely related to the amidohydrolase MscE which is found in both clusters. Its detailed function remains elusive at the current stage of research.

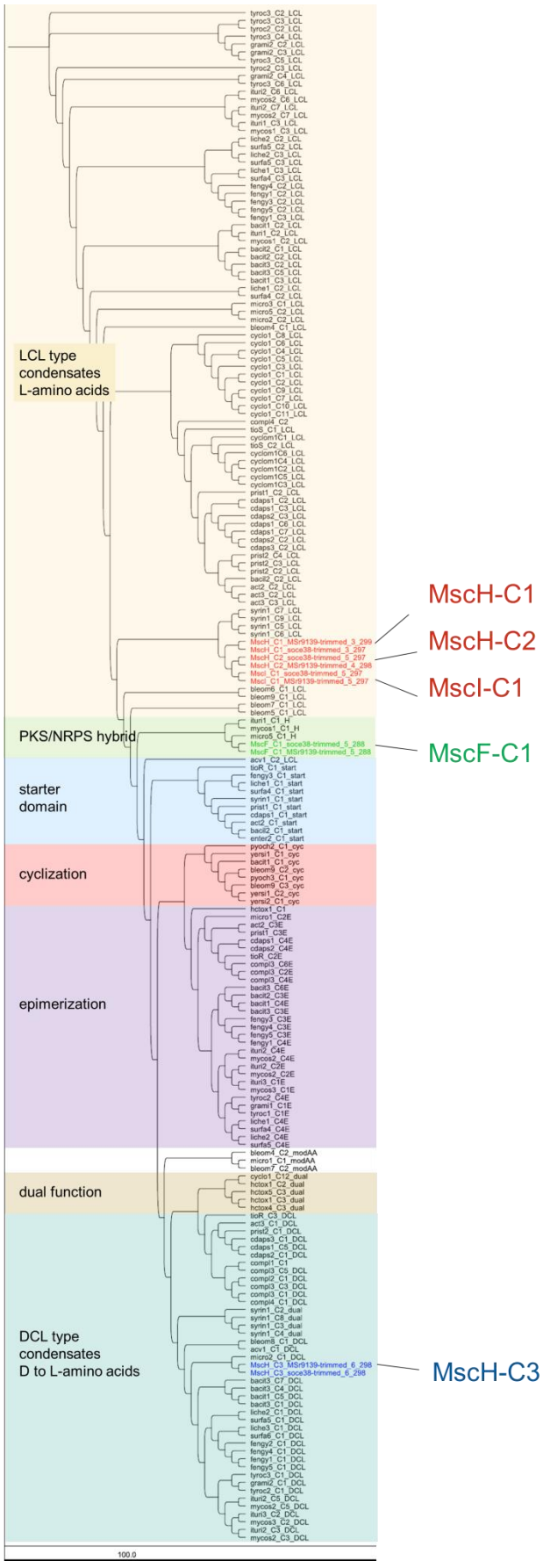


Figure S12: Phylogenetic tree highlighting different C domain types and the C-domains found in the msc biosynthetic gene cluster in So ce38 and MSr9139.

MT-domains

The *N*-methyltransferase present in module 1 of MscH is found in both clusters and is responsible for the *N*-methylation of the glycine right next to the pyrrolidone ring. Both of them harbor the common consensus-motif GxGxG at the *N*-terminal end which is responsible for SAM Co-factor binding (Table S13).[8] These *N*-methyltransferases belong to the UbiE/COQ5 family. In addition, the *msc*-cluster of *Jahnella* sp. encodes a second methyltransferase domain which is located in MscA. This methyl transferase has homology to the PRMT5 family. However, the MT domain is likely inactive as the SAM binding motif GxGxG is changed to GxGxE. The exchange of the small glycine to a charged glutamic acid may prevent cofactor binding.

Table S13: Active site analysis of the methyltransferase domains (MT) within the msc *biosynthetic gene cluster.*

active consensus	G	x	G	X	G
MscA-MT$_{\text{Jahnella sp.}}$	G	T	G	K	E
MscH-MT$_{\text{So ce38}}$	G	C	G	T	G
MscH-MT$_{\text{Jahnella sp.}}$	G	C	G	T	G

Determination of configuration

The conformation of all stereogenic centers of the isolated microsclerodermins is identical to that of the microsclerodermins C – I and confirms the assumed stereochemistry of the pedeins.[9–11] In detail, the scaffold is based on an *R*-tryptophan, a (*3R*)-4-NH_2-3-OH-butyric acid moiety, a side chain with (*2S,3R,4S,5S*)-configuration and an (*R,R*)-pyrrolidone.

Absolute configuration of 3-aminobutyric acid and tryptophan

Microsclerodermin has several stereogenic centers. We identified the configuration of the γ-aminobutyric acid (GABA) moiety and tryptophan by acid hydrolysis followed by advanced Marfey analysis using FDLA reagent. (3*S*)-GABA was ordered from Sigma and used as a reference together with *S*-tryptophan and *S*-6-chloro-tryptophan. Derivatization of the γ-amino functionality with L-FDLA and D-FDLA results in two products with reproducible retention time differences (3 seconds using 10 technical replicates, see Figure S13) which allows annotation of the stereochemistry. In conclusion, the new microsclerodermins L and M have an (3*R)*-configured carbon atom in the GABA subunit which is in agreement with the other microsclerodermins reported.

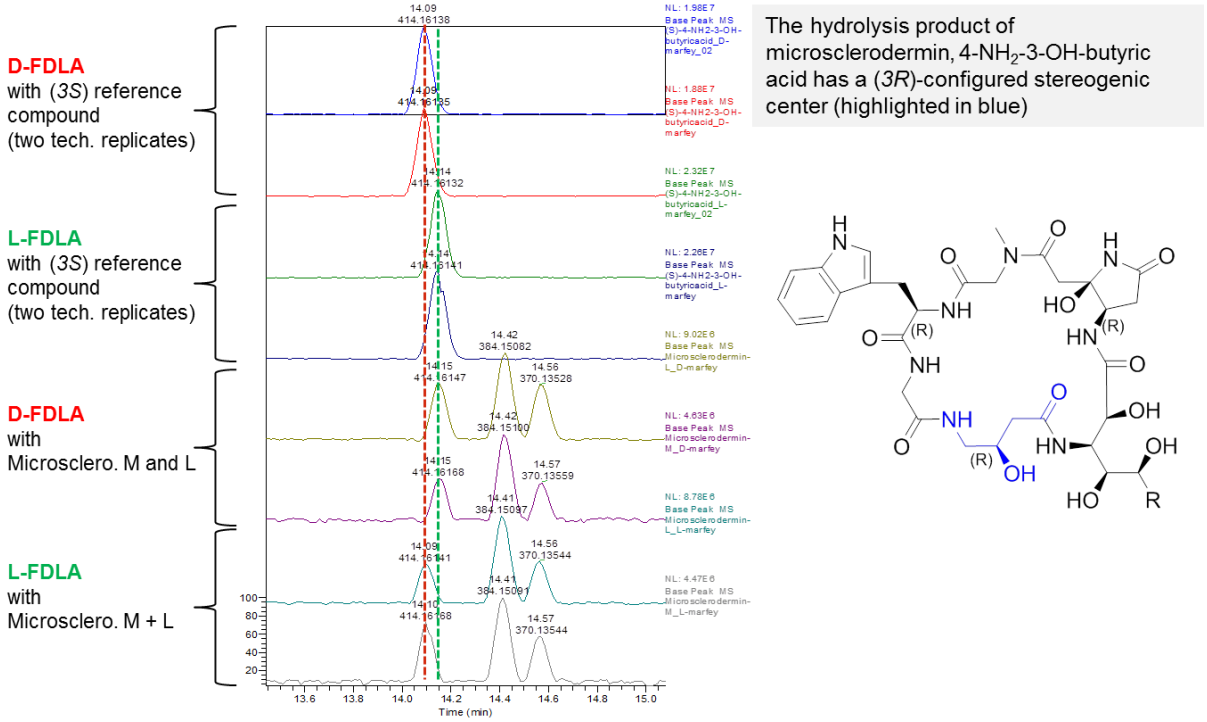


Figure S13: Overlay of LC-MS chromatograms highlighting the retention time difference of the GABA subunit when derivatized with Marfey's reagent (FDLA). The graph features the (3S)-GABA reference compound as well as microsclerodermin M and L.

The tryptophan in microsclerodermin M, D and L is *R*-configured. This applies to all microsclerodermins known so far and is in agreement with the epimerization domain found within the respective module of the biosynthetic assembly lines in So ce38 and MSr9139. Microsclerodermin M was compared to *S*-tryptophan whereas microsclerodermin D and L were compared to a *S*-6-Cl-tryptophan reference (Figure S14).

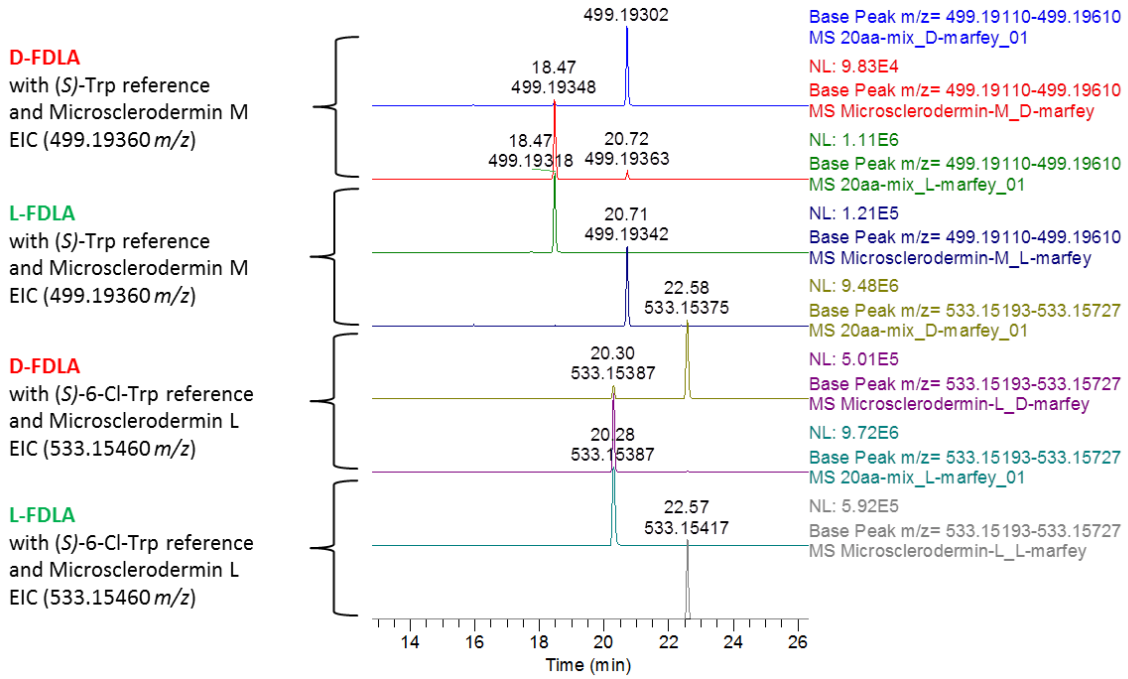


Figure S14: Overlay of the LC-MS chromatograms for tryptophan characterization using advanced Marfey's reagent (FDLA). The stereogenic center of tryptophan is (R)-configured in microsclerodermin M and L.

Configuration of the pyrrolidone moiety

The pyrrolidone moiety of the microsclerodermins is derived from an asparagine as indicated by feeding labeled *S*-asparagine. Incorporation of labeled *S*-asparagine suggested that the pyrrolidone ring keeps this configuration. However, this turned out to be wrong. Microsclerodermins M, D and L underwent the same chemical degradation steps as reported for the microsclerodermins A-I.[9,10,12] After dehydration at the pyrrolidone moiety the respective dehydromicrosclerodermins were subjected to ozonolysis. The reaction products were hydrolyzed and derivatized with Marfey's reagent. Marfey derivatization and HPLC-MS measurements revealed an *R*-configured pyrrolidone moiety in the microsclerodermin derivatives M, L, and D. Although this chemical degradation protocol is used for the microsclerodermins A – I in literature, the method is not that elegant in terms of the result as ozonolysis, oxidative workup and hydrolysis of *R*-tryptophan will result in *R*-asparate as well. At the same time this procedure will generate aspartate out of the pyrrolidone moiety. By doing a short hydrolysis step of not more than 60 minutes we observe a notable amount of *R*-asparagine in the sample which can only originate from the pyrrolidone moiety. Based on these results we were able to identify *R*-tryptophan and *R*-pyrrolidone in microsclerodermin M and D. We verified the results by applying the above mentioned protocol to pure *R*-tryptophan as well as to the synthesized, enantiomeric pure *S*-dehydropyrrolidone fragment (98 % ee). *R*-tryptophan is converted to *R*-aspartate while the *S*-dehydropyrrolidone fragment is converted to asparagine and aspartate with a *S*:*R*-ratio of approximately 4:1. This partial inversion clarifies the origin of the *R*- and *S*-asparagine mixture which is observed in all samples of microsclerodermin, there with the opposite *S*:*R*-ratio of 1:4 (Figures S15)

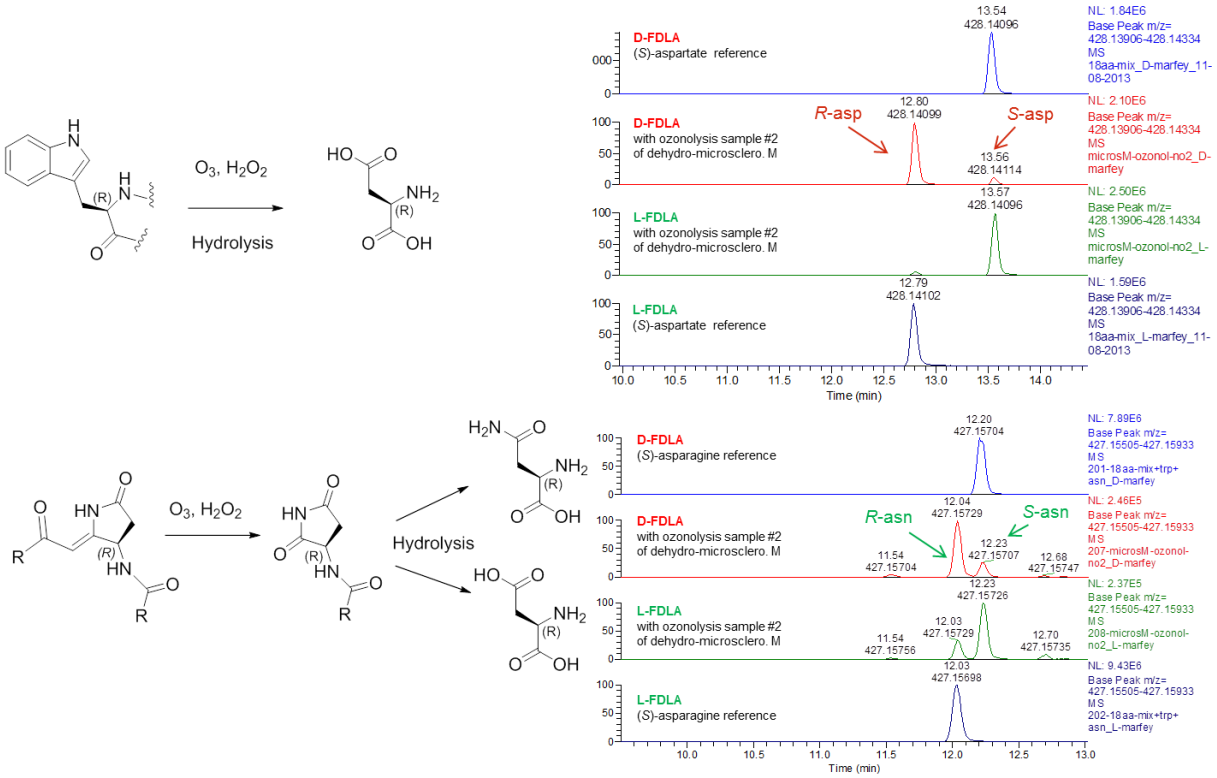


Figure S15: Marfey analysis result of ozonolysis product of microsclerodermin M. The chromatograms show the peaks for asparagine and aspartate, both as an overlay display with the respective S-configured reference compound.

Microsclerodermin D and L are both isolated from strain MSr9139. In case of microsclerodermin L, the pyrrolidone moiety has an additional methoxy group resulting in a 3-methoxy-aspartate after the above mentioned reactions whereas microsclerodermin D has the same pyrrolidone moiety as the M derivative (So ce38). Unfortunately we were not able to get suitable enantiomeric pure reference compounds for 3-methoxy-aspartate. However, based on the rule of thumb that an L-Marfey derivative of an L-configured amino acid elutes earlier than that of an D-Marfey derivative[13] we can conclude that it is the other way round in our sample and we do have an D-configured amino acid, (*2R*)-3-methoxy-aspartate, present in microsclerodermin L (Figure S16). This is supported by the detection of *R*-asparagine in the microsclerodermin D derived sample which is isolated from the same strain and thereby underlying the same biosynthesis. The stereogenic center at the methoxy group was not assigned by this method.

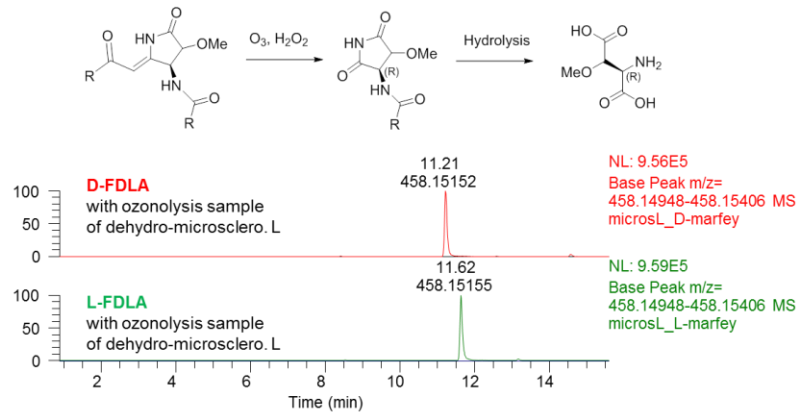


Figure S 16: Marfey analysis result of ozonolysis product of microsclerodermin L. The chromatograms show the peaks for 3-MeO-aspartate. As the D-marfey derivative (upper part) elutes before the L-marfey derivative (lower part) the 3-MeO-aspartate has a 2R-configured stereogenic center. There is no reference compound available.

Configuration of the side chain

The side chain's vicinal hydroxyl groups OH-4 and OH-5 were converted to an acetonide in order to identify the relative stereochemistry of this part of the molecule. The acetonide formation was done for microsclerodermin M from *S. cellulosum* So ce38 and microsclerodermin L from *Jahnella* sp. MSr9139. Acetonides were purified using the same HPLC method as for the natural product isolation. Purity was checked by HPLC-MS analysis prior to NMR analysis. Respective NOE couplings are detected using selective irradiation experiments for the protons attached to the acetonide moiety. The NOE couplings are identical to those determined for the microsclerodermins known to literature. Based on this result we conclude to have the same relative stereochemistry as reported for microsclerodermins A – I which have a *2S,3R,4S,5S*-configured side chain. We assigned the absolute stereochemistry of the side chain by diol cleavage at OH-4 and OH-5. The product is oxidized to the respective carboxylic acid and hydrolyzed using 6 N HCl. After Marfey derivatization we observed a peak for L-threo-ß-OH-aspartate (*2S,3S* configuration) which would be in agreement with the *2S,3R,4S,5S*-configured side chain.

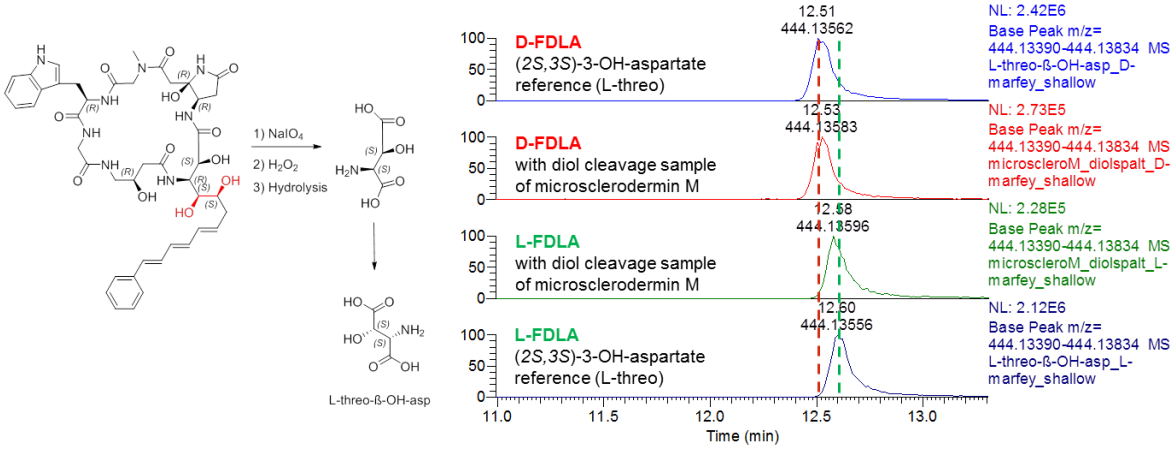


Figure S17: Marfey analysis of diol cleavage product (Microsclerodermin M).Peak identified by reference compound L-threo-ß-OH-aspartate.

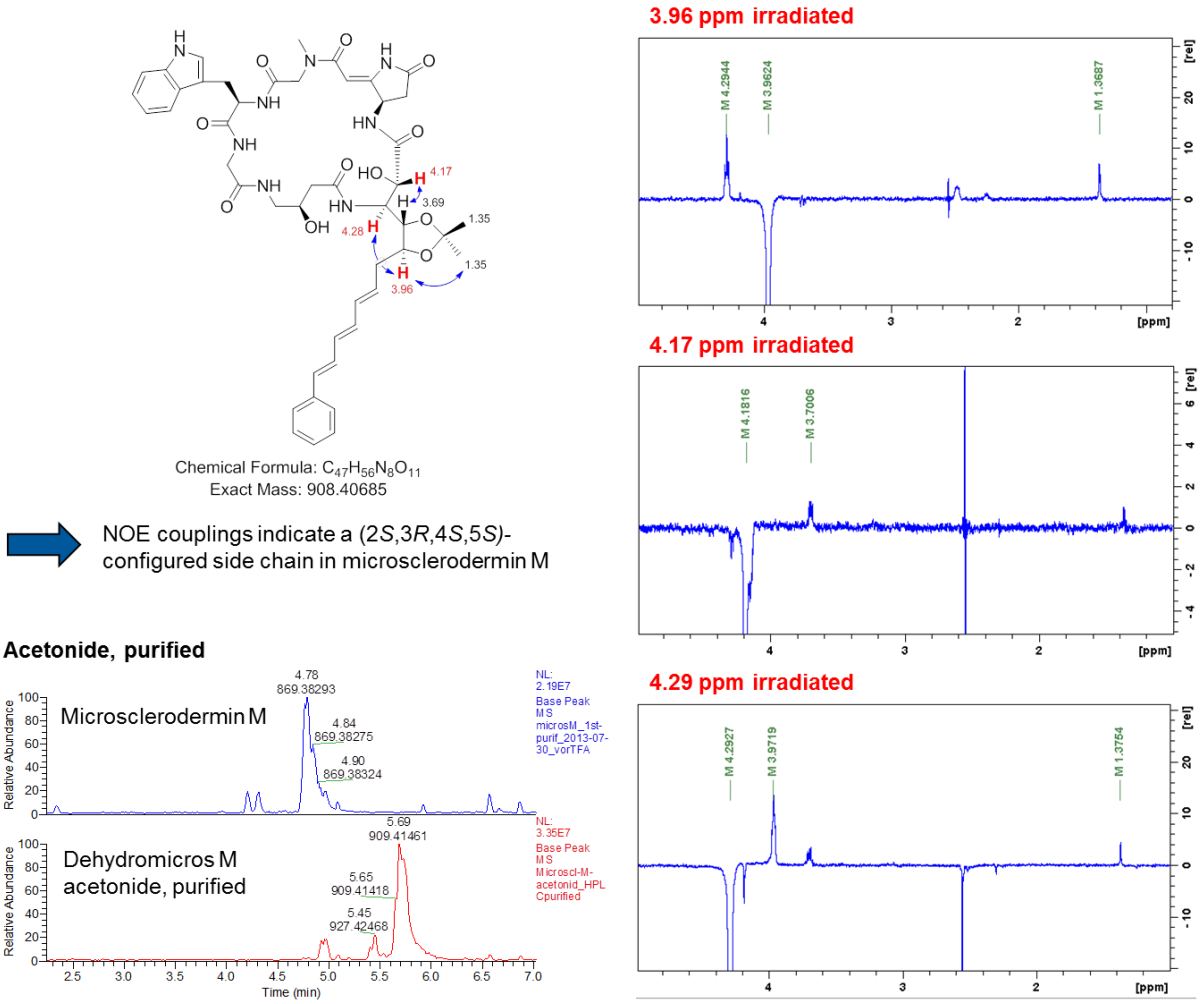


Figure S18: Acetonide purification and selective irradiation experiments of dehydromicrosclerodermin M.

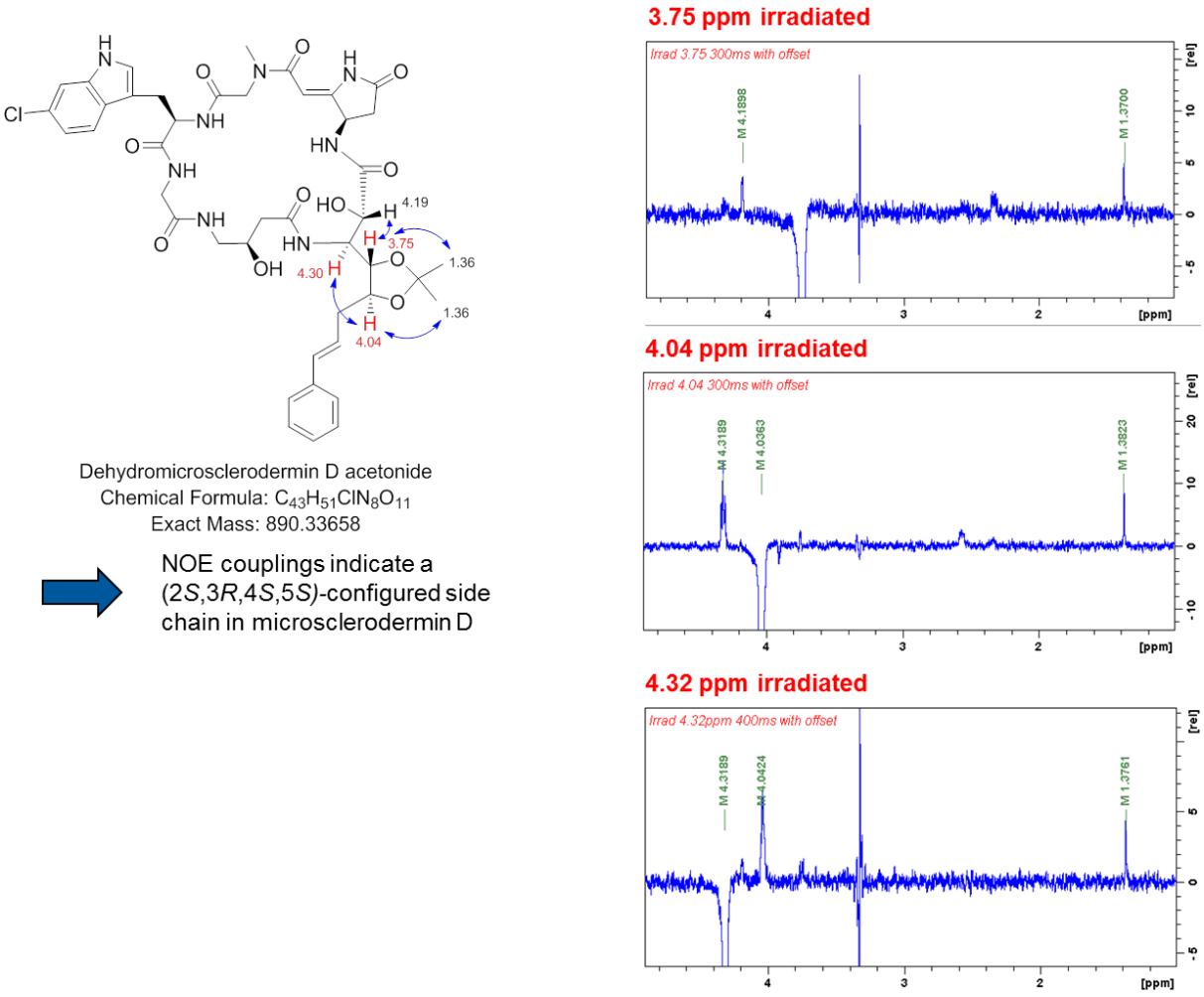


Figure S19: Acetonide purification and selective irradiation experiments of dehydromicrosclerodermin D.

Experimental section

Dehydration of microsclerodermin. A solution of microsclerodermin (2 mg) in 100 µl DMSO was diluted with 1 % aqueous TFA (400 µl). The mixture was kept at 40 °C for 15 minutes and dried afterwards using a GeneVac Evaporator (GeneVac Ltd, Suffolk, UK). Quantitative conversion to the respective dehydromicrosclerodermin was verified by dissolving the products in methanol and measuring them by LC-MS analysis.

Ozonolysis of dehydromicrosclerodermin. Dehydromicrosclerodermin (100 µg) was dissolved in 400 µL dry methanol. A stream of ozone was bubbled through the cooled solution (-78 °C) for 20 minutes. The excess reagent was removed by a stream of nitrogen before warming the solution to 20 °C. Remaining solvent was then removed under reduced pressure and the dry ozonide dissolved in 1 mL formic acid/hydrogen peroxide (2:1, v/v). The reaction mixture was heated to 80 °C for 30 min before drying the mixture under reduced pressure. The sample was then used for hydrolysis and Marfey derivatization.

Diol cleavage of microsclerodermin. Microsclerodermin (500 µg) was dissolved in 50 µL DMSO and added to 300 µL of sodium periodate solution in water (0.5 mg/mL), adjusted to pH 4.0 with acetic acid. The sample was stirred at room temperature for 16 hours and dried afterwards. The residue was dissolved in hydrogen peroxide (1 mL) and formic acid (0.5 mL)

and heated to 70 °C for 20 min. After removing the solvent under reduced pressure the sample was ready for hydrolysis and Marfey derivatization.

Acetonide of dehydromicrosclerodermin. Under an atmosphere of dry nitrogen 200 µL of 2,2-dimethoxy propane is added to a stirred solution of dehydromicrosclerodermin (2 mg) in 150 µL dry DMF. The reaction is initiated by adding catalytic amounts of pyridinium p-toluene sulfonic acid and stirred for 16 hours at room temperature. The reaction is quenched by adding 20 µl pyridine and dried using a stream of nitrogen. The residues are diluted in 100 µL H_2O/DMSO (1:1, v/v) and the respective acetonide purified using the same HPLC method as for the initial purification of the microsclerodermins in this work.

Marfey derivatization protocol. Marfey derivatization was performed with NMR, ozonolysis and diol-cleavage samples of the compounds. The protocol is as follows:

- Put at least 50 µg of sample into a 1.4 mL glass vial
- Add 100 µL of 6 N HCl. Fill the vial with nitrogen and close it. Keep it at 110 °C for 45 minutes. Open the vial to let dry at 110 °C for another 15 minutes. Do not exceed 60 min as tryptophan easily decomposes.
- Dissolve the residues in 110 µL H_2O. Prepare two 1.5 mL PP tubes and add 50 µL of the aqueous solution in each.
- Add 20 µL of 1 N $NaHCO_3$ in each tube (pH adjusted to approx. 9)
- Add 20 µL of 1 % (w/v) Marfey's reagent in acetone: D-FDLA or L-FDLA, respectively.
- Keep for 1 hour at 40 °C, 700 rpm.
- Add 10 µL of 2 N HCl to stop reaction and 300 µL of ACN to end up with 400 µL total volume.
- Centrifuge the sample and measure the supernatant by HPLC-MS

HPLC-MS analysis of Marfey samples. All measurements were performed on a Dionex Ultimate 3000 RSLC system using a Waters BEH C18, 100 x 2.1 mm, 1.7 µm d_p column by injection of 1 µl sample. Separation was achieved by a gradient using (A) H_2O + 0.1 % FA to (B) ACN + 0.1 % FA at a flow rate of 550 µl/min and 45 °C. The gradient was as follows: starting at 5 % B to increase to 10 % B in 1 min, from 1 to 15 min increase to 35 % B, from 15 to 22 min increase to 50 % B, from 22 to 25 min increase to 80 % B. After a 1 min hold at 80 % B the system was reequilibrated with initial conditions for 5 minutes. UV data was acquired at 340 ± 8 nm and MS detection was performed simultaneously. Coupling the HPLC to the MS was supported by an Advion Triversa Nanomate nano-ESI system attached to a Thermo Fisher Orbitrap. LC flow is split to 500 nL/min before entering the ion source. Mass spectra were acquired in centroid mode ranging from 150 – 1000 m/z at a resolution of R = 30000.

HPLC-MS based screening

We searched our in-house database for producers of microsclerodermins. This screening included LC-MS data from 791 myxobacterial extracts covering a broad diversity of myxobacterial genera (Figure S20). Eleven different *Sorangium* sp. were identified as producers of microsclerodermin M. Microsclerodermin M is a rather instable compound that readily forms isobaric isomers. This effect is most likely related to the unsaturated side chain and causes several peaks that nearly coelute under standard LC-MS screening conditions. Thus, extracted ion chromatograms (EICs) for the $[M\text{-}H_2O+H]^+$ signal at 869.38283 *m/z* (blue trace, Figure S21) show up with a broadened peak including multiple isomers. In addition, the $[M+H]^+$ signal at 887.39340 *m/z* is detected albeit with lower abundance. We observe an additional, later eluting peak in the EICs with a $[M+H]^+$ signal at 869.38283 *m/z*, most likely caused by a loss of water at the pyrrolidone ring upon extract preparation and storage.

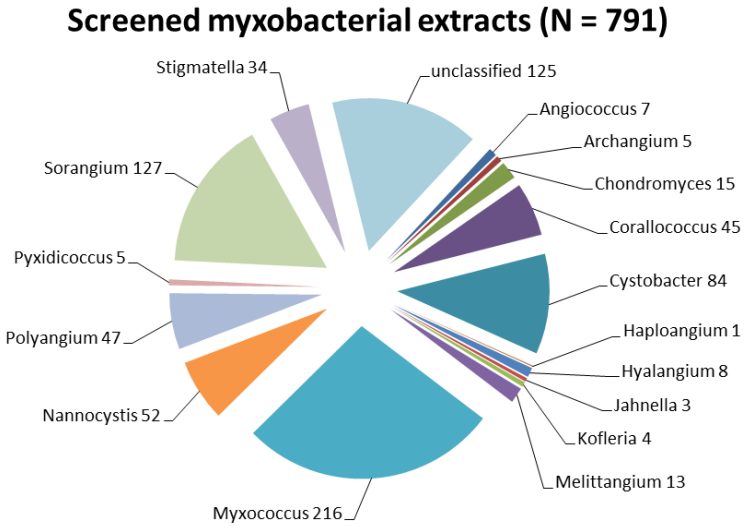


Figure S20: Myxobacterial genera as covered in our LC-MS based screening data set.

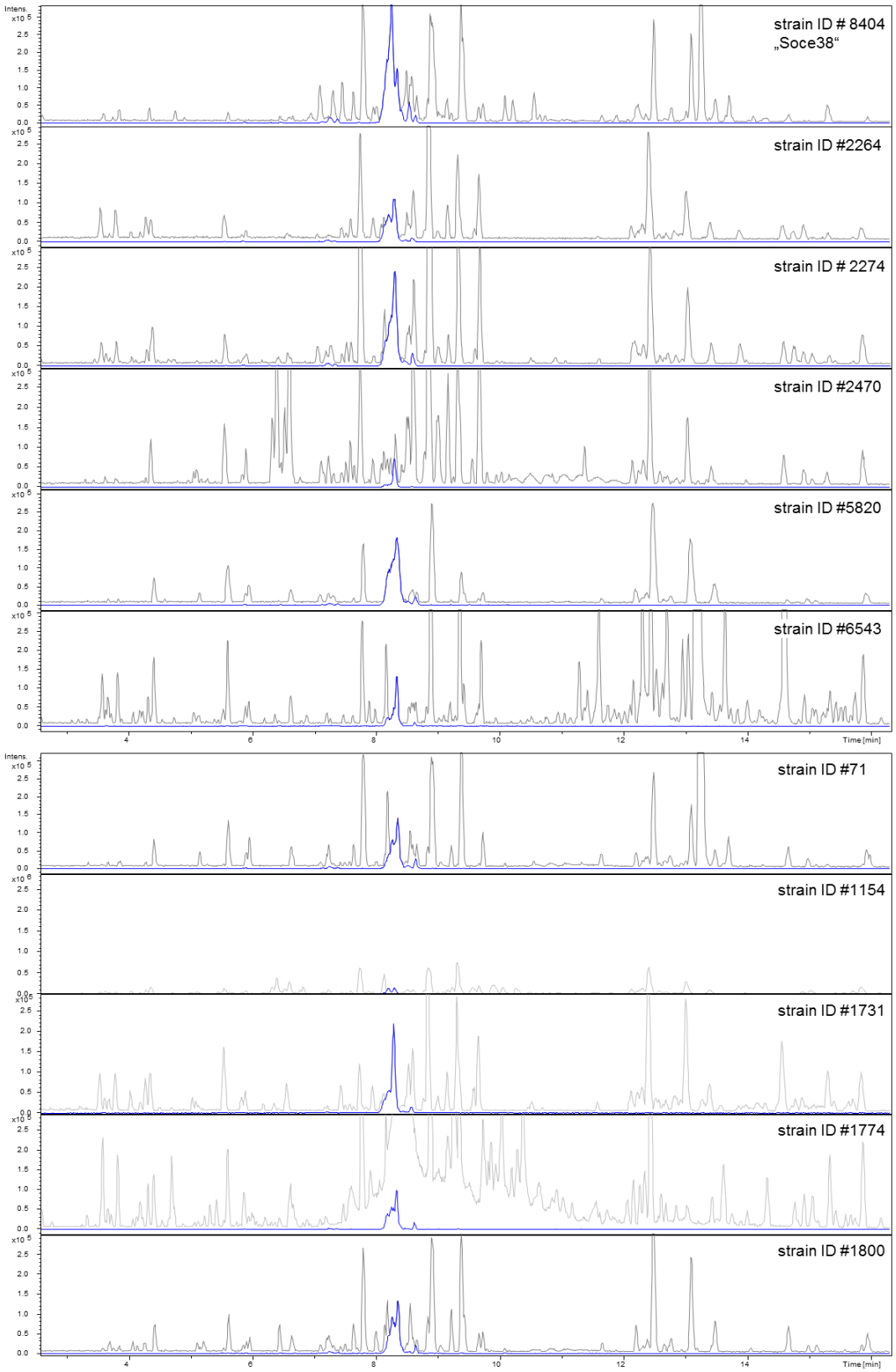


Figure S21: List of microsclerodermin M producers as identified by searching our in-house database. All strains are of the genus Sorangium*. Microsclerodermin M and its isobaric isomers are highlighted by the EIC of 869.38283 m/z ($[M\text{-}H_2O+H]^+$ ion, blue trace). Abundance of derivatives is related to the different producer strain. Irregular peak shape is due to isobaric, almost coeluting derivatives occurring during storage of the extracts. Identification of microsclerodermin M is based on accurate* m/z*, isotope pattern fit, retention time and MS^2 pattern compared to the reference compound.*

Another group of myxobacteria is producing the known microsclerodermin D as well as pedein A and B. In addition, one new derivative (Microsclerodermin L) is identified in this work (Figure S22). The strain *Chondromyces pediculatus* Cm p3 is also part of this group.[11]

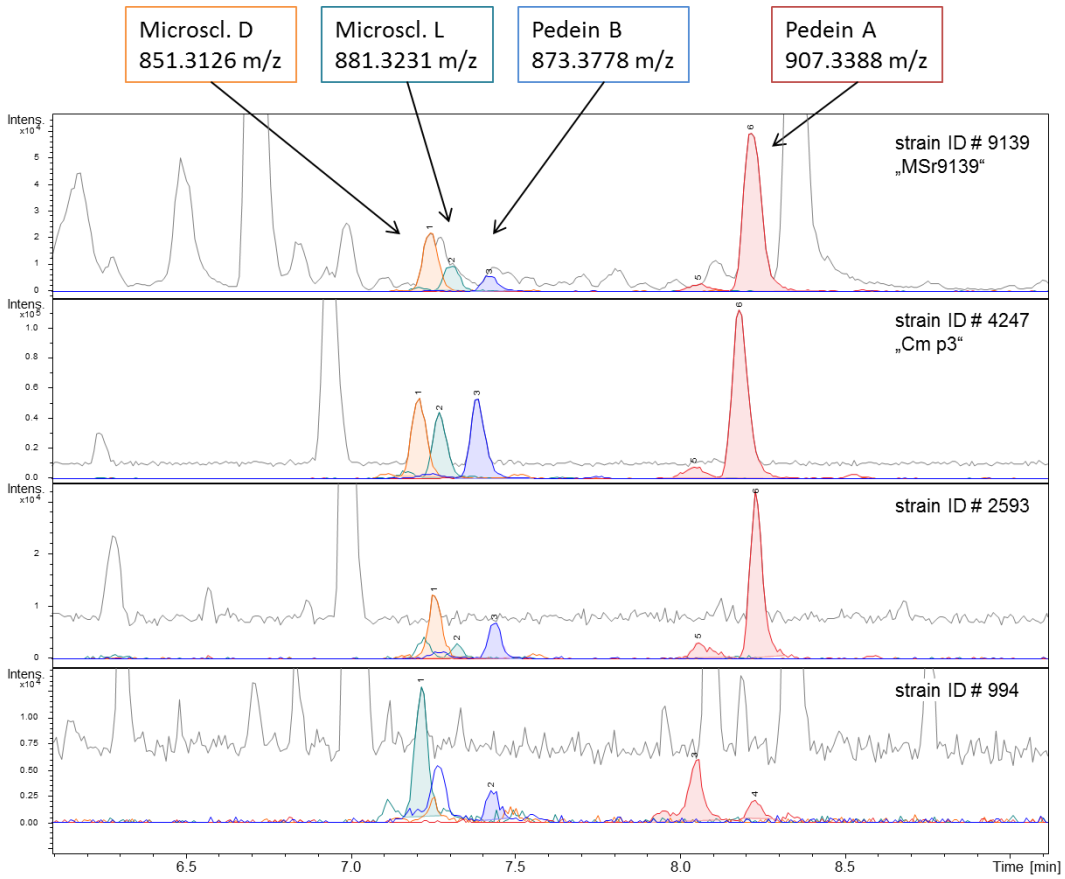


Figure S22: The known derivatives of microsclerodermin in four producer strains are highlighted by colored extracted ion chromatograms (EICs) overlaying the base peak chromatogram (BPC, grey). The producers are of the genus Jahnella *and* Chondromyces*. Abundance of derivatives is related to the different producer strain. Identification of microsclerodermin derivatives is based on accurate* m/z*, isotope pattern fit, retention time and MS² pattern compared to the reference compounds.*

LC-MS method for screening. All measurements were performed on a Dionex Ultimate 3000 RSLC system using a BEH C18, 100 x 2.1 mm, 1.7 µm dp column (Waters, Germany). Separation of 1 µl sample was achieved by a linear gradient from (A) H2O + 0.1 % FA to (B) ACN + 0.1 % FA at a flow rate of 600 µl/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5 % B, followed by an increase to 95 % B in 18 min to end up with a 2 min step at 95 % B before reequilibration with initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75 µl/min before entering the maXis 4G hr-ToF mass spectrometer (Bruker Daltonics, Germany) using the Apollo ESI source. Mass spectra were acquired in centroid mode ranging from 150 – 2500 m/z at 2 Hz scan rate.

Double Bond rearrangement

A retro-biosynthetic approach of the microsclerodermin side chain will initially result in a putative incorporation of a very uncommon C3 extender unit. This finding is owing to the double bond order which usually indicates the C2 extender units in between the double bonds. However, in microsclerodermin this does not hold true as the double bonds can rearrange in order to form a π-system conjugated to the phenyl ring. Taking this into account, the retro-biosynthetic proposal is based on a phenylacetate derived starter unit followed by C2 units, which is in agreement with conventional biosynthetic logic and the results from feeding experiments with labeled precursors. It remains elusive if this rearrangement is the result of sample extraction and workup or already takes place during/after biosynthesis *in vivo*.

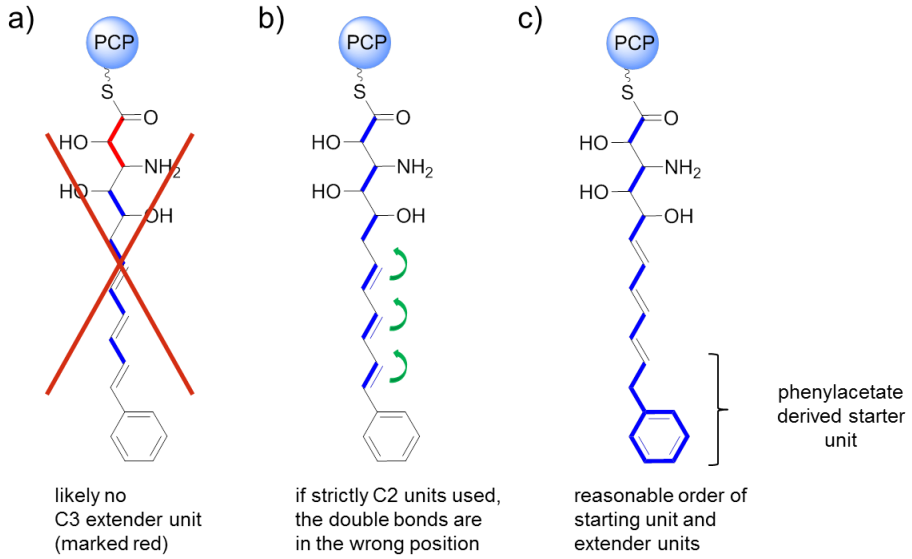
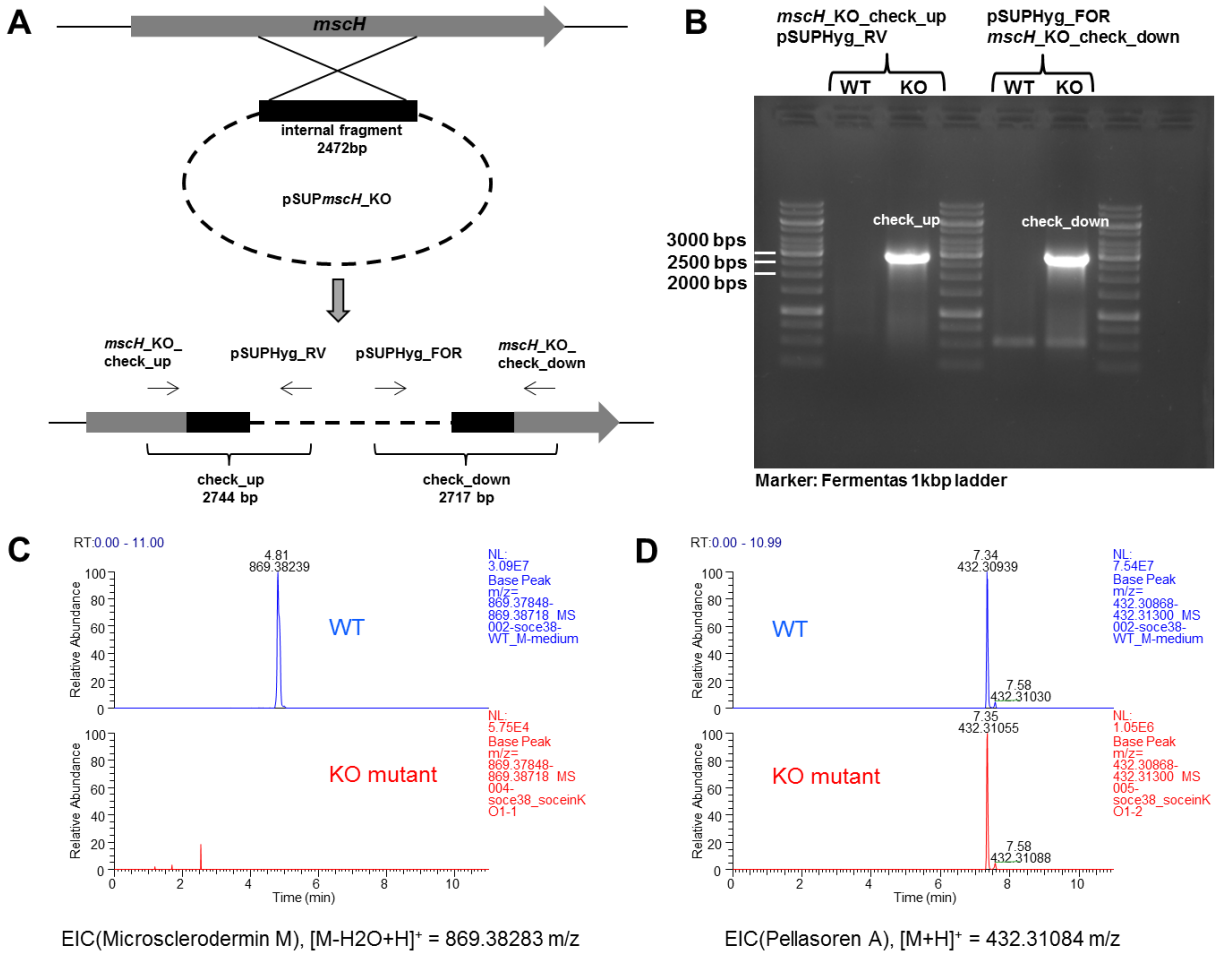


Figure S23: Retro-biosynthetic proposal for the microsclerodermin side chain. a) C3 extender unit is likely not incorporated during biosynthesis. b) A double bond rearrangement leads to C2 extender units. c) A reasonable biosynthetic proposal is starting with a phenylacetate starter followed by several C2 extender units.

Targeted inactivation of the *msc* locus in So ce38

Genetic modification of the So ce38 strain was accomplished according to a previously described protocol.[14] Putative single cross-over transformants were grown in H-Medium (0.2 % soybean meal, 0.2 % glucose, 0.8% starch, 0.2% yeast-extract, 0.1 % $CaCl_2 \cdot 2\ H_2O$, 0.1 % $MgSO_4 \cdot 7H_2O$, 8 mg/L FeEDTA, 50 mM HEPES, adjusted to pH 7.4 using 10N KOH) supplemented with hygromycin (100 µg/mL) and 1 % adsorber resin (XAD-16, Rohm & Haas) for at least 5 days at 180 rpm and 30 °C. Methanolic extracts of cells and the XAD resin were analyzed by LC-MS and compared with WT extracts. For isolation of chromosomal DNA the mutant strain was grown in M-medium containing 100 µg/mL hygromycin. The integration of the plasmid at the right position in the genome was verified by PCR as shown in supporting Figure S24, using mutant and wild-type DNA as templates.



EIC(Microsclerodermin M), $[M-H2O+H]^+$ = 869.38283 m/z

EIC(Pellasoren A), $[M+H]^+$ = 432.31084 m/z

Figure S24: (A) Site of homologous recombination in So ce38. (B) Verification of KO-mutant So ce38::pSUPmscH_KO using PCR resulting in two amplicons "check_up" and "check_down". (C) Abolishment of microsclerodermin production as detected by comparative LC-MS analysis of wildtype and mutant strains. An extracted ion chromatogram for the most abundant $[M-H_2O+H]^+$ signal is shown. (D) Regular secondary metabolite production of the mutant strain was verified by identification of a known secondary metabolite in both wild-type and mutant strain.

Table S14: Primers used for amplification and verification of the microsclerodermin knock-out mutant in So ce38.

oligonucleotide	sequence
*mscH*_KO_for	GAT CCA GCG CTG GTT CCT CG
*mscH*_KO_for	ACT CGC CCT CGC GGA GGT TCT
pSUPHyg_fwd	ATG TAG CAC CTG AAG TCA GCC
pSUPHyg_rev	ACG CAT ATA GCG CTA GCA GC
*mscH*_KO_check_up	ACA ACT TCT TCG CGC TCG G
*mscH*_KO_check_down	TCG TCG TAC GAG AGC CGG

Feeding experiments using labeled precursors

Feeding in *S. cellulosum* So ce38

$^{15}N_2$-$^{13}C_4$-L-asparagine
(completely incorporated)
$\Delta m_{theor.}$ = 2 x ^{15}N and 4 x ^{13}C
= (2 x 0.997035 Da) + (4 x 1.003355 Da)
= 6.00749 Da
$\Delta m/z_{meas.}$ = 6.0075

^{15}N-$^{13}C_9$-L-phenylalanine:
(eight carbons incorparated)
8 x ^{13}C
$\Delta m_{theor.}$ = 8 x ^{13}C
= 8 x 1.003355 Da
= 8.02684 Da
$\Delta m/z_{meas.}$ = 8.0267

$^{13}C_6$-L-phenylalanine (ring label)
(six carbons incorporated)
6 x ^{13}C
$\Delta m_{theor.}$ = 6 x ^{13}C
= 6 x 1.003355 Da
= 6.02013 Da
$\Delta m/z_{meas.}$ = 6.02023

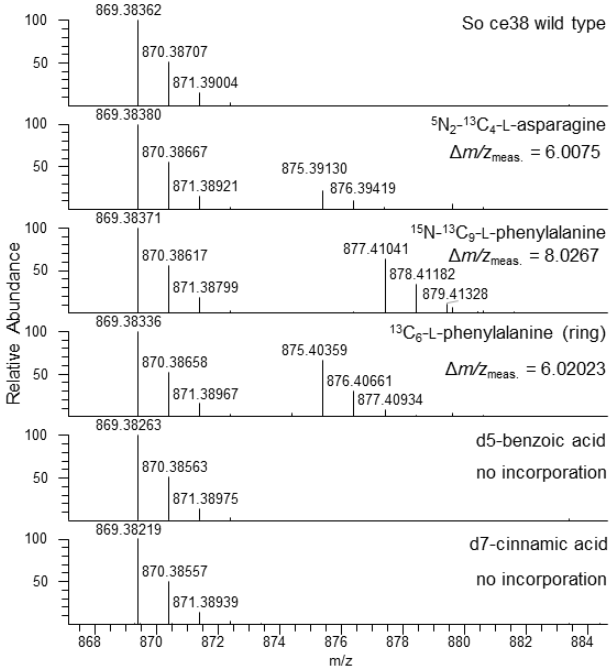


Figure S25: Isotopic peak pattern of the highly abundant [M-H2O+H]+ signal at 869.38283 m/z proves incorporation of labeled precursors into microsclerodermin M (So ce38). The observed mass shifts do perfectly fit to the heavy isotopes that were incorporated.

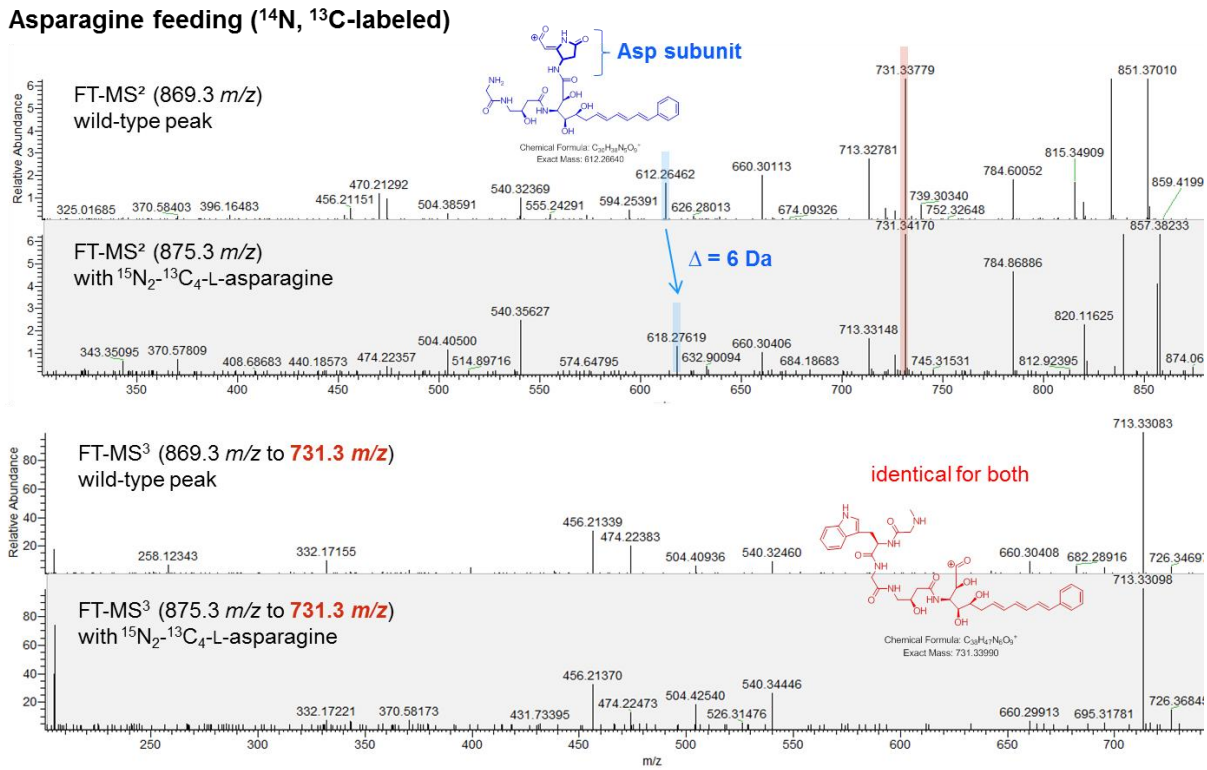


Figure S26: Incorporation of labeled L-asparagine subunit characterized by MS² and MS³ spectra acquisition. Data shows the feeding experiment in S.cellulosum So ce38. The most abundant fragment with 731.3 m/z is identical for the wild type peak and the enriched derivative (see MS³ data).

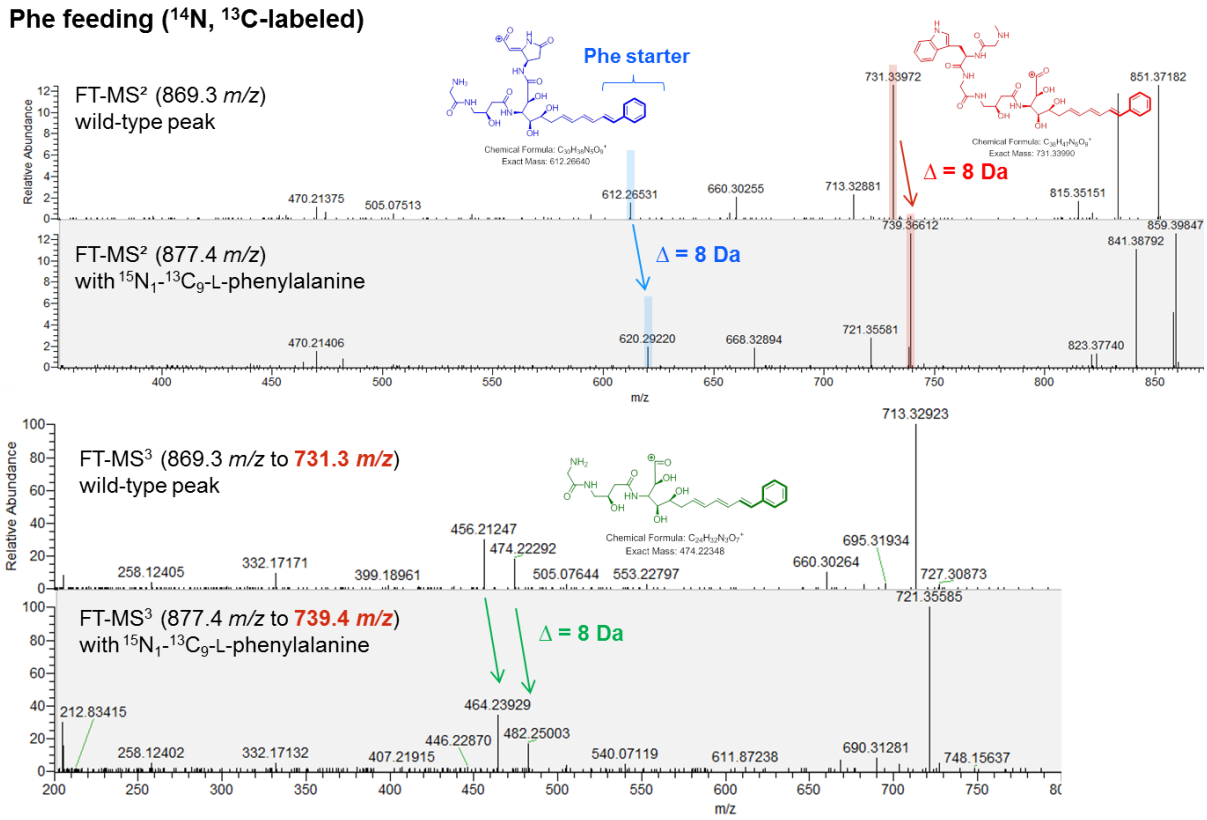


Figure S27: Incorporation of labeled phenylalanine subunit characterized by MS² and MS³ spectra acquisition. Data shows the feeding experiment in S.cellulosum So ce38. The mass difference of 8 Da is present in all the highlighted fragments indicating that phenylalanine is incorporated into the side chain.

Feeding in *Jahnella* sp. MSr9139

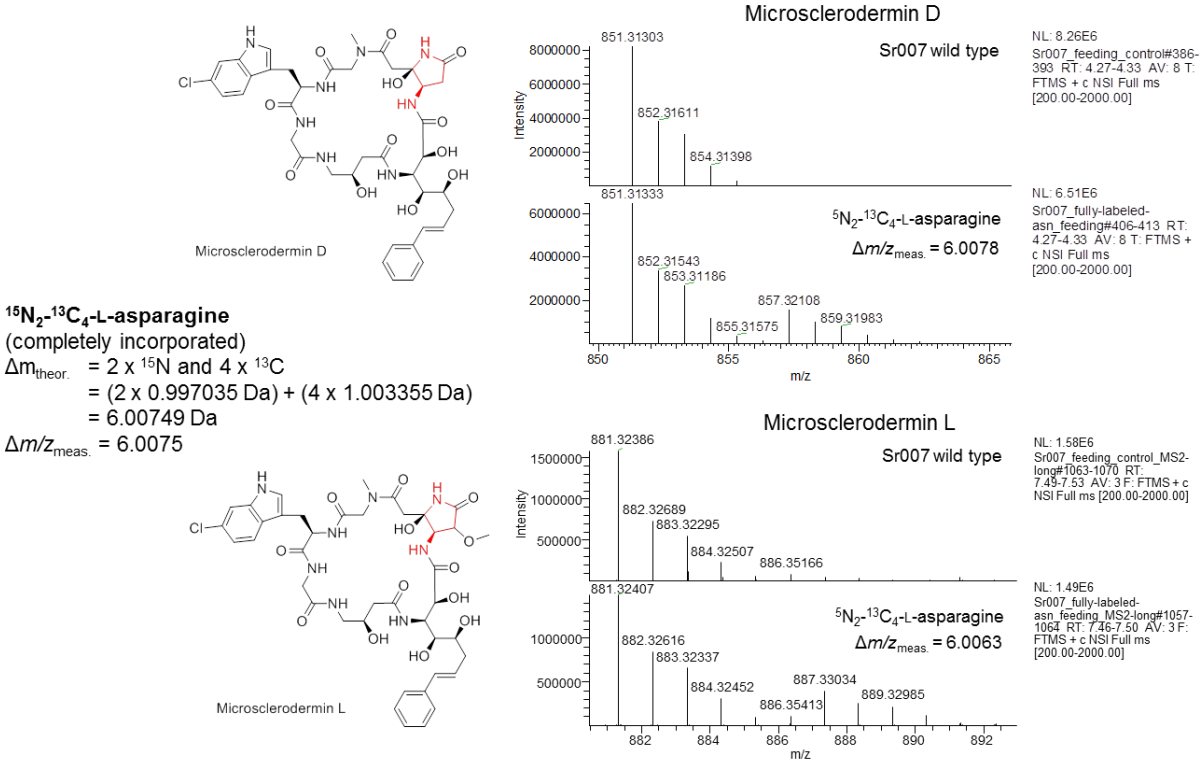


Figure S28: Isotopic peak pattern of the highly abundant [M-H_2O+H]$^+$ signals for microsclerodermin D (88 m/z) and L (881.32312 m/z) prove incorporation of labeled precursor in Jahnella sp. MSr9139. The observed mass shifts do perfectly fit to the heavy isotopes that were incorporated.

Experimental section

Feeding in So ce38. Cultivation of So ce38 was performed in 25 mL H medium using 1 % (w/v) XAD-16. Labeled amino acids (5 mg) were dissolved in 250 mM HCl (320 µL) and sterile filtered. The labeled precursor stock (0.16 mg/mL) was added in three portions to the culture. 80 µL were added on first and third day followed by the remaining 160 µL on the fifth day of cultivation. Labeled benzoic acid and cinammic acid were dissolved in 80 µL DMSO and added in 2 x 20 µL and 40 µl portions to the respective culture. The cultures were harvested after 9 days and extracted with 2 x 25 mL methanol. The organic solvent was removed and the residuals dissolved in 500 µL methanol.

Feeding in MSr9139. Cultivation of MSr9139 was performed in 25 mL VY/2 medium without addition of XAD-16. Labeled asparagine (4.5 mg) was dissolved in 250 mM HCl (320 µL) and sterile filtered. The labeled precursor stock (0.14 mg/mL) was added in three portions to the culture. . 80 µL were added on first and second day followed by the remaining 160 µL on the fourth day of cultivation. The culture was harvested after 6 days and extracted with 2 x 25 mL methanol. The organic solvent was removed and the residuals dissolved in 500 µL methanol.

LC-MS analysis of feeding experiments. Measurements were performed on a Dionex Ultimate 3000 RSLC system using a Waters BEH C18, 50 x 2.1 mm, 1.7 µm d_p column by injection of two µl methanolic sample. Separation was achieved by a linear gradient with (A) H_2O + 0.1 % FA to (B) ACN + 0.1 % FA at a flow rate of 600 µL/min and 45 °C. The gradient

was initiated by a 0.3 min isocratic step at 5 % B, followed by an increase to 95 % B in 9 min to end up with a 1 min flush step at 95 % B before reequilibration with initial conditions. UV and MS detection were performed simultaneously. Coupling the HPLC to the MS was supported by an Advion Triversa Nanomate nano-ESI system attached to a Thermo Fisher Orbitrap. Mass spectra were acquired in centroid mode ranging from 200 – 2000 *m/z* at a resolution of R = 30000. The measurements were repeated to obtain high resolution CID data in order to support the experiment with annotation of the obtained fragments.

Bioactivity Testing

Bacterial Cultures. All microorganisms were handled under standard conditions recommended by the depositor. Overnight cultures of microorganisms were prepared in EBS medium (0.5 % peptone casein, 0.5 % proteose peptone, 0.1 % peptone meat, 0.1 % yeast extract; pH 7.0) or TSB medium (1.7 % peptone casein, 0.3 % peptone soymeal, 0.25 % glucose, 0.5 % NaCl, 0.25 % K_2HPO_4; pH 7.3). The latter medium was used for *E. faecalis* and *S. pneumonia* cultures. Yeast and fungi were grown in Myc medium (1% phytone peptone, 1 % glucose, 50 mM HEPES, pH 7.0).

Microbial Susceptibility Assay (MIC). Overnight cultures of microorganisms were diluted to OD_{600} 0.01 (bacteria) or 0.05 (yeast/fungi) in the respective growth medium. Serial dilutions of compounds were prepared as duplicates in sterile 96-well plates. The cell suspension was added and microorganisms were grown overnight on a microplate shaker (750 rpm, 30 °C or 37 °C). Growth inhibition was assessed by measuring the OD_{600} on a plate reader. MIC_{50} values were determined as average relative to respective control samples by sigmoidal curve fitting.

Reference List

(1) Delcher, A. L.; Bratke, K. A.; Powers, E. C.; Salzberg, S. L. *Bioinformatics* **2007**, *23*, 673–9.
(2) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res.* **2011**, *39*, W339–46.
(3) Yadav, G.; Gokhale, R. S.; Mohanty, D. *Journal of Molecular Biology* **2003**, *328*, 335–363.
(4) Ziemert, N.; Podell, S.; Penn, K.; Badger, J. H.; Allen, E.; Jensen, P. R. *PLoS ONE* **2012**, *7*, e34064.
(5) Caffrey, P. *ChemBioChem* **2003**, *4*, 654–7.
(6) Kwan, D. H.; Tosin, M.; Schläger, N.; Schulz, F.; Leadlay, P. F. *Org. Biomol. Chem.* **2011**, *9*, 2053–6.
(7) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res.* **2011**, *39*, W362–7.
(8) Kozbial, P. Z.; Mushegian, A. R. *BMC Struct. Biol.* **2005**, *5*, 19.
(9) Schmidt, E. W.; John Faulkner, D. *Tetrahedron* **1998**, *54*, 3043–3056.
(10) Qureshi, A.; Colin, P. L.; Faulkner, D. J. *Tetrahedron* **2000**, *56*, 3679–3685.
(11) Kunze, B.; Böhlendorf, B.; Reichenbach, H.; Höfle, G. *J. Antibiot.* **2008**, *61*, 18–26.
(12) Bewley, C. A.; Debitus, C.; Faulkner, D. J. *J. Am. Chem. Soc.* **1994**, *116*, 7631–7636.
(13) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Analytical Chemistry* **1997**, *69*, 5146–5151.
(14) Kopp, M.; Irschik, H.; Gross, F.; Perlova, O.; Sandmann, A.; Gerth, K.; Müller, R. *J. Biotechnol.* **2004**, *107*, 29–40.

NMR spectra collection

NMR Data – Microsclerodermin D

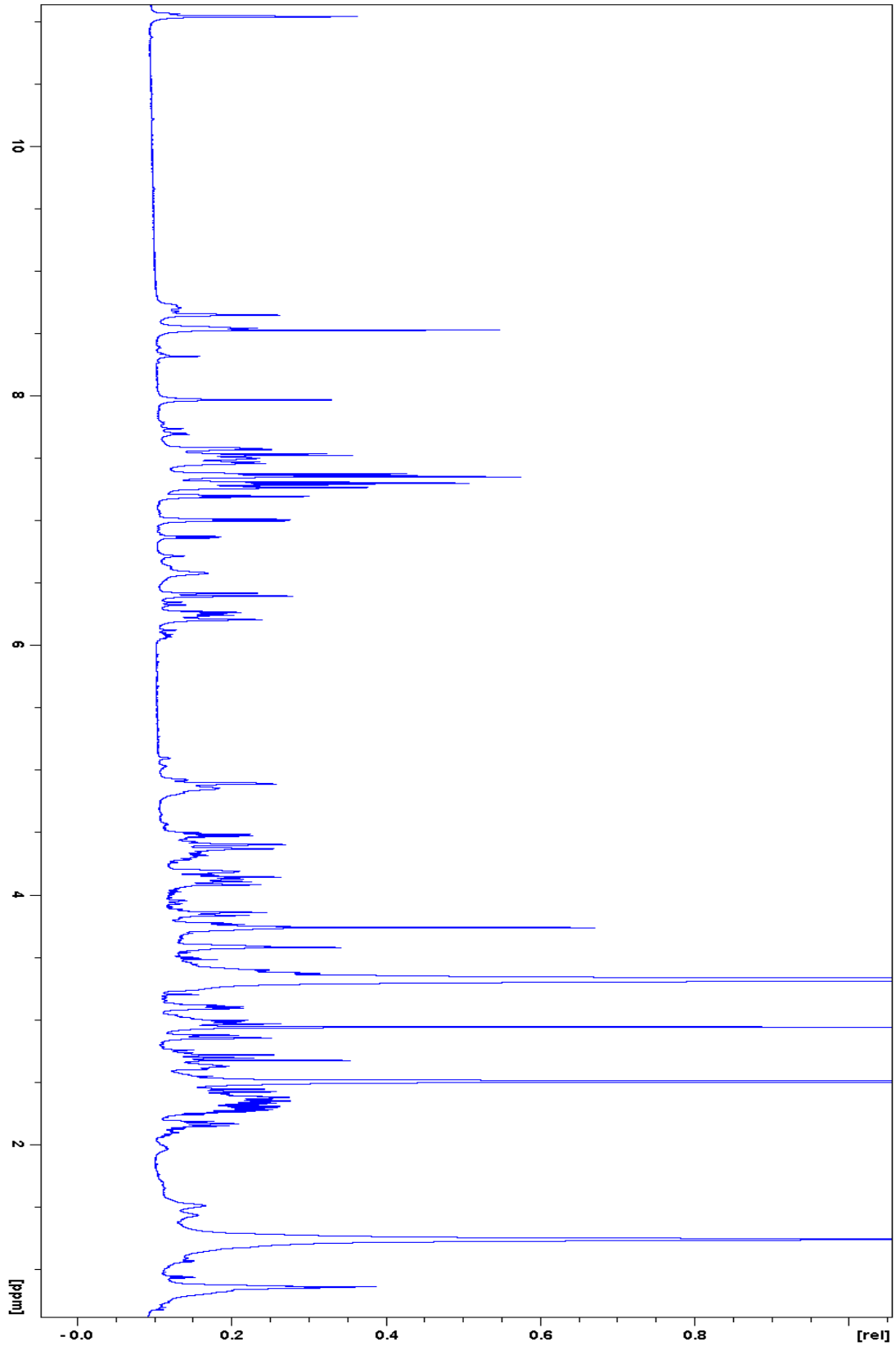


Figure 29: ^{1}H NMR spectrum of microsclerodermin D in DMSO-d_6

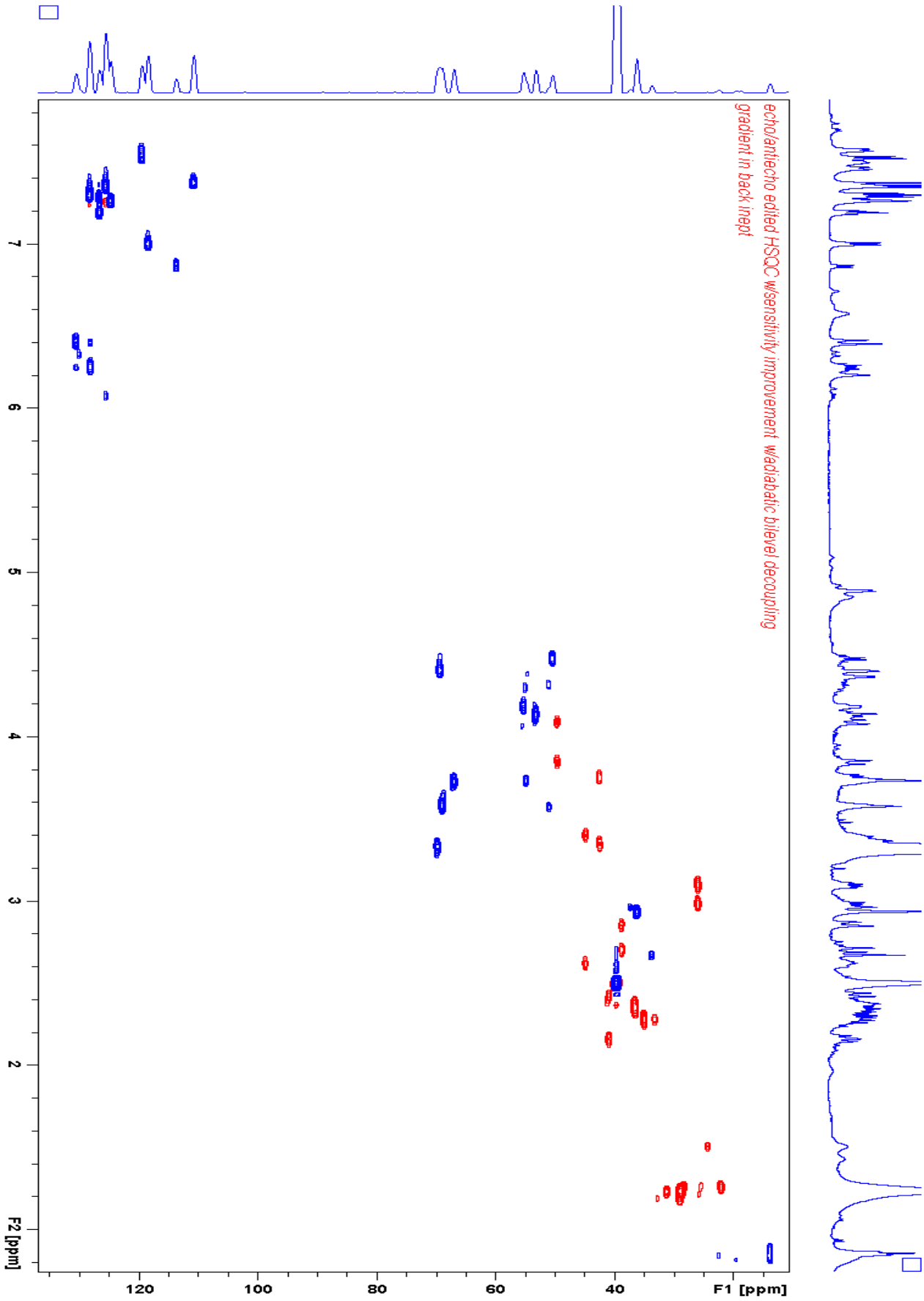


Figure 30: HSQC spectrum of microsclerodermin D in DMSO-d_6

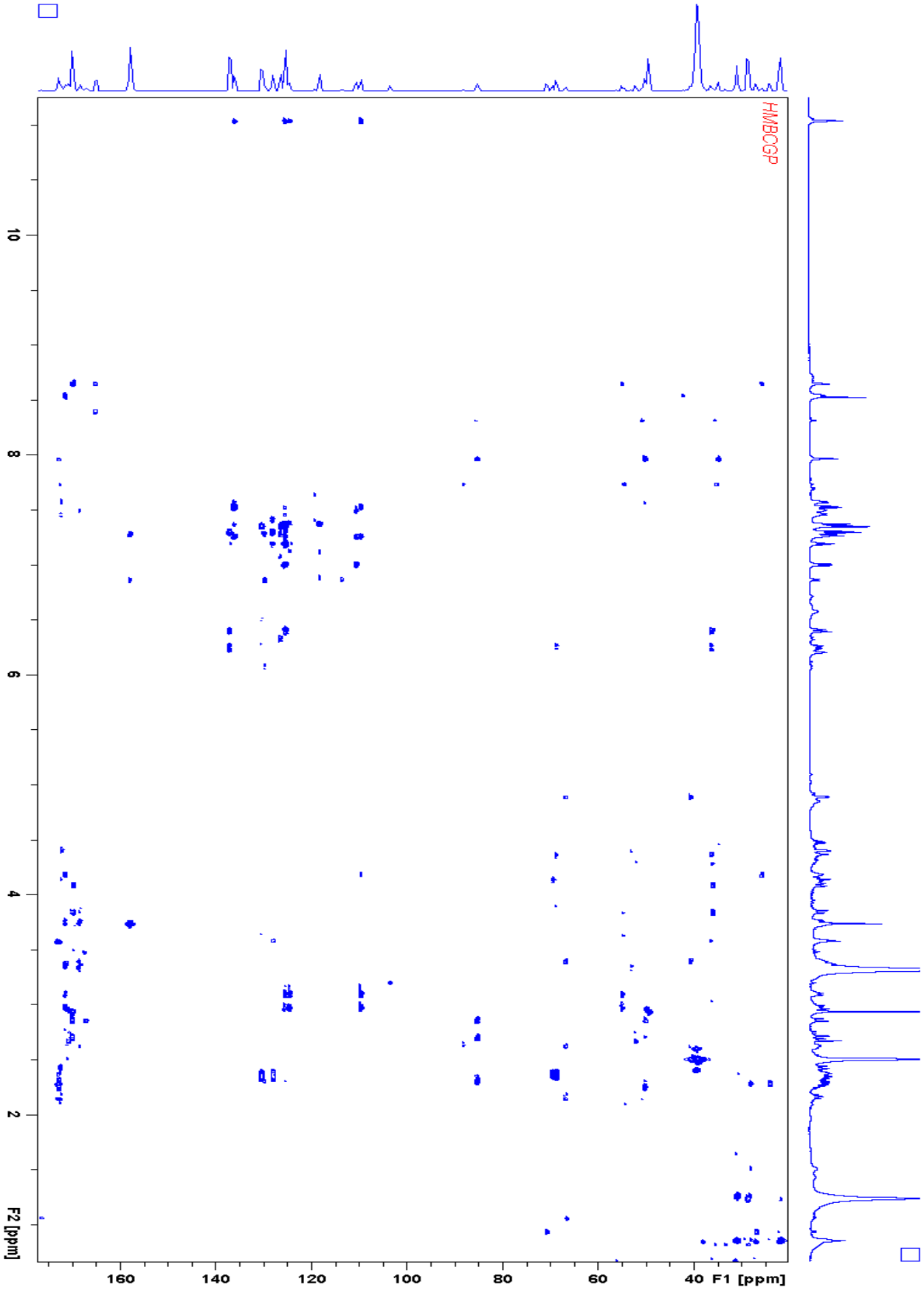


Figure 31: HMBC spectrum of microsclerodermin D in DMSO-d_6

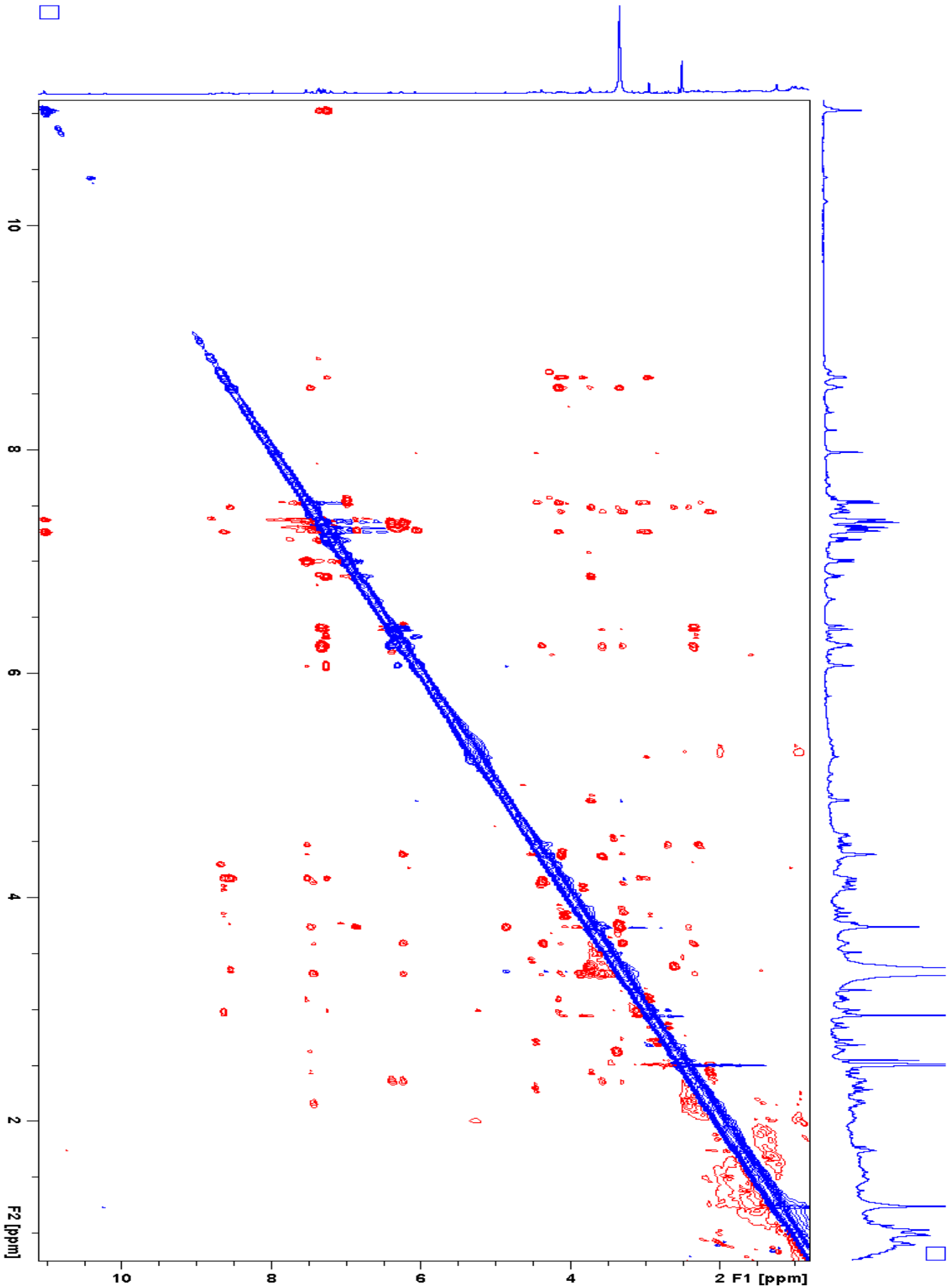


Figure 32: ROESY spectrum of microsclerodermin D in DMSO-d_6

NMR Data – Dehydromicrosclerodermin D acetonide

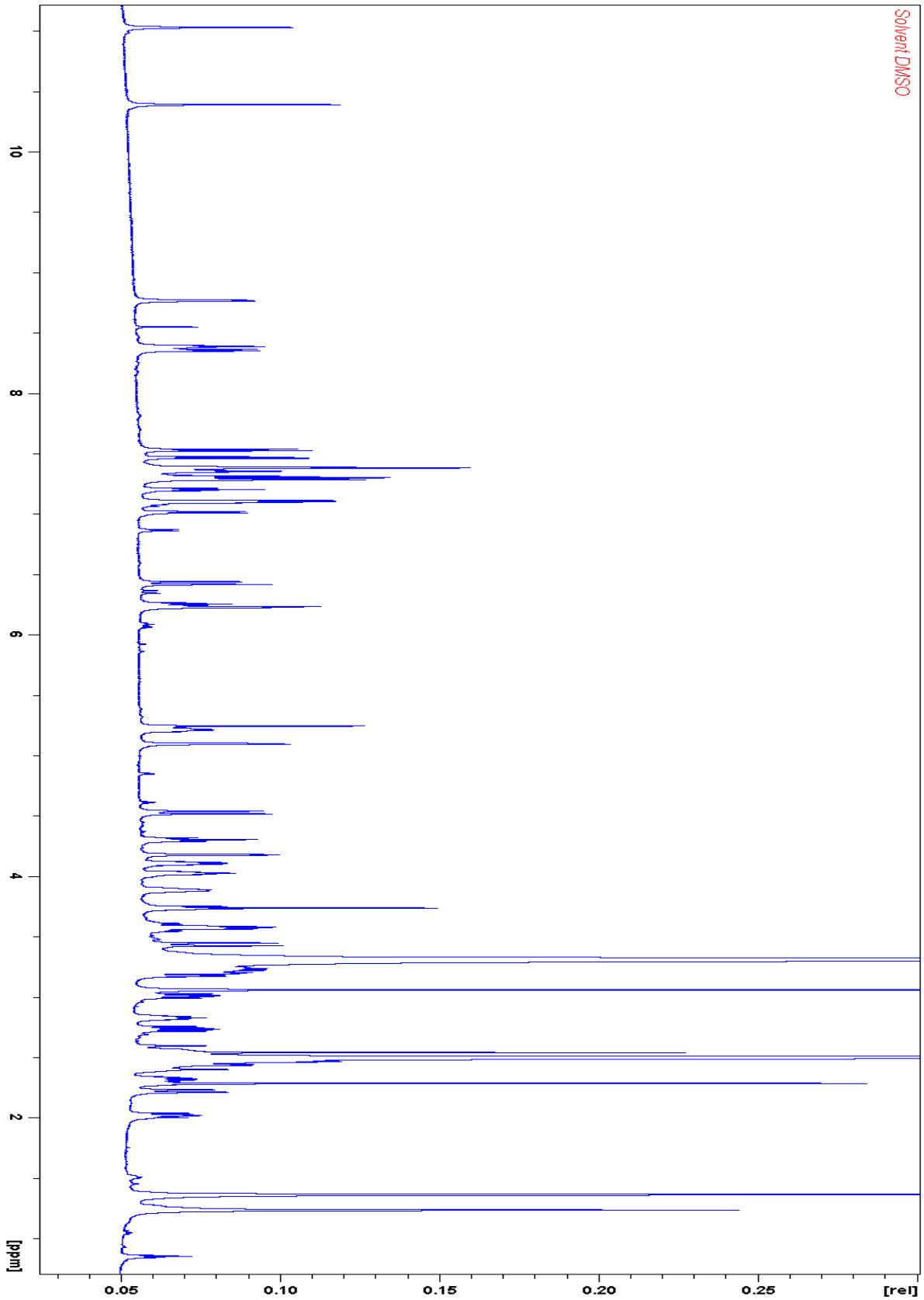


Figure 33: ^{1}H NMR spectrum of dehydromicrosclerodermin D - acetonide in DMSO-d_6

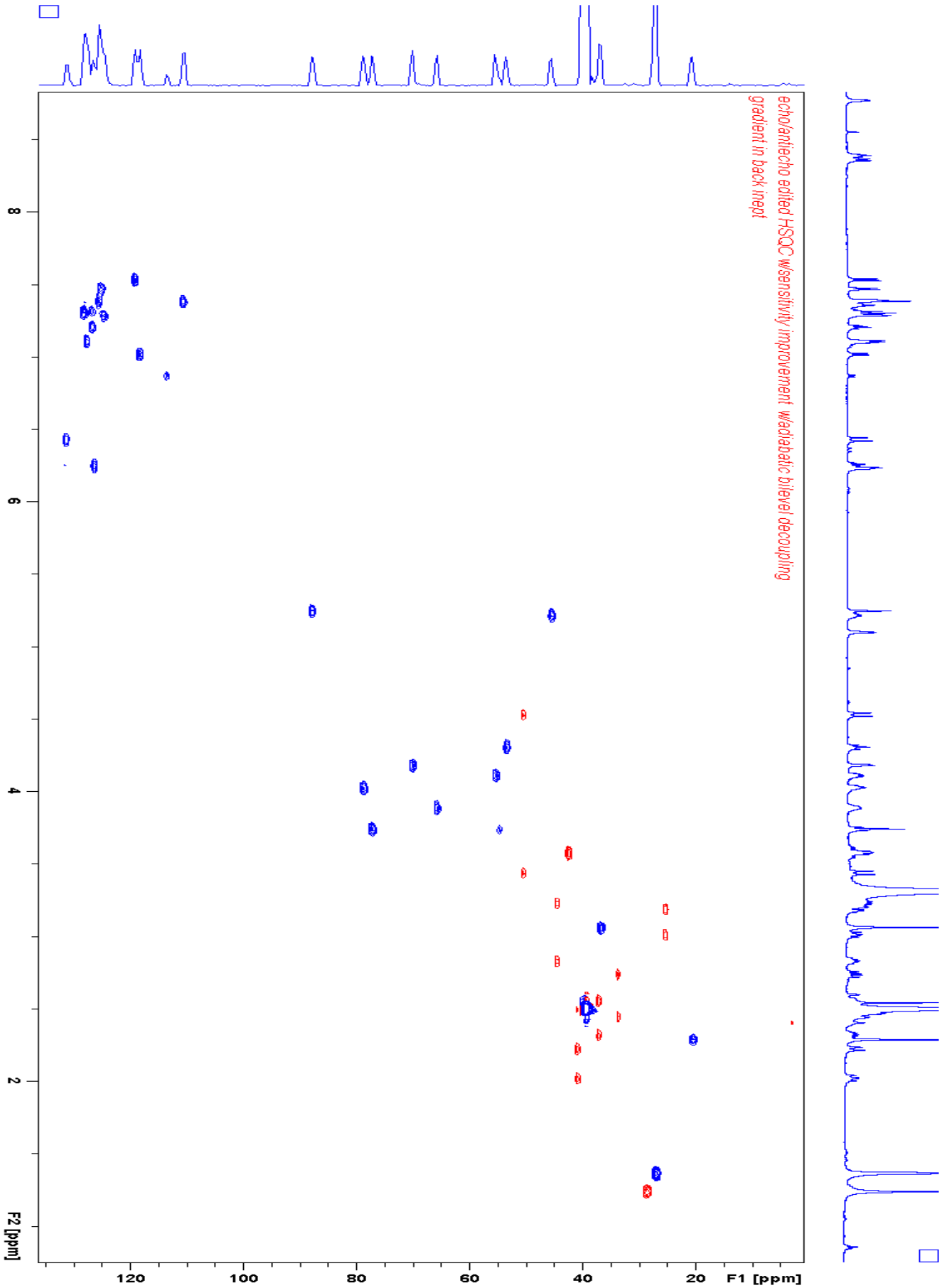


Figure 34: HSQC spectrum of dehydromicrosclerodermin D - acetonide in DMSO-d_6

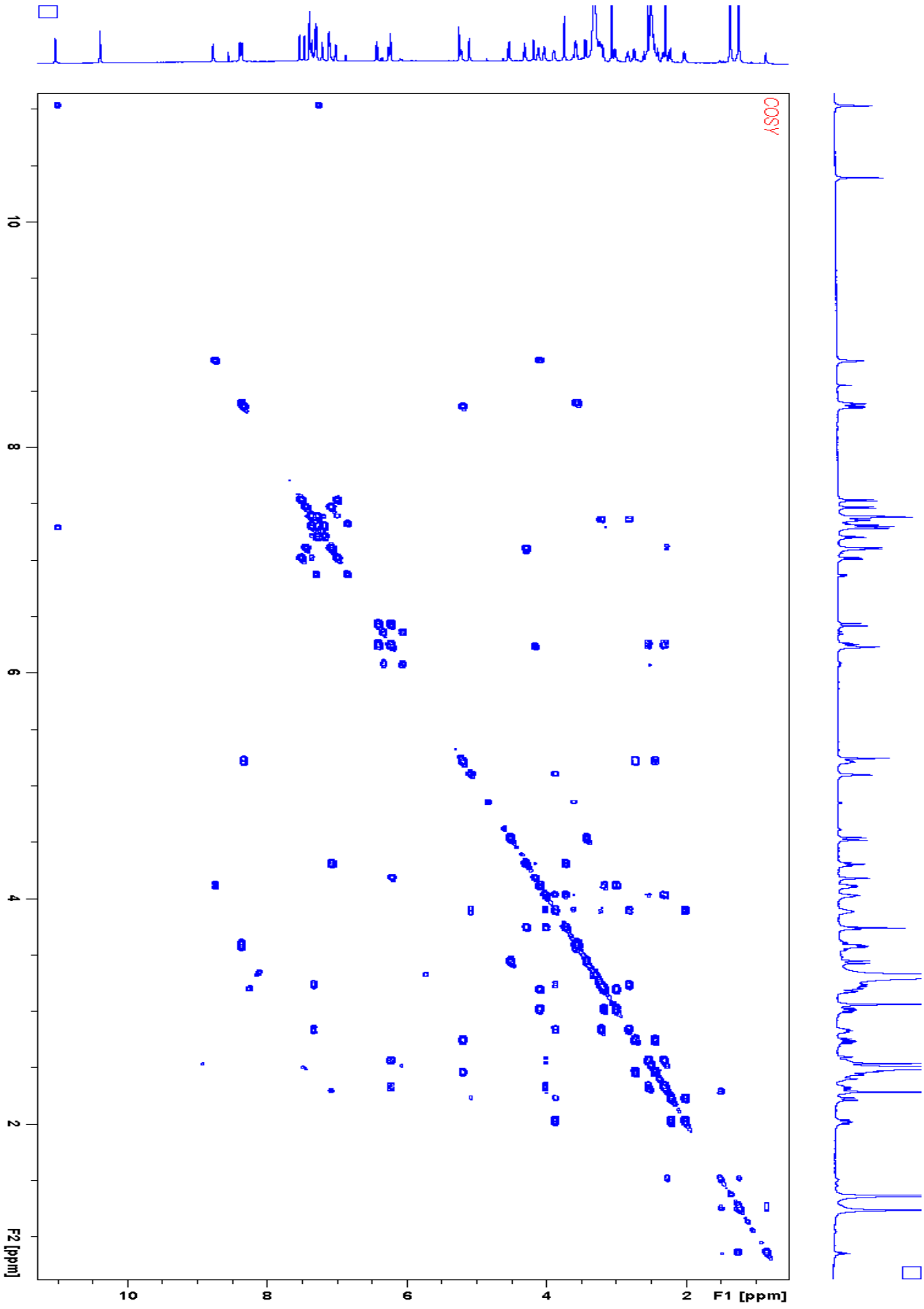


Figure 35: COSY spectrum of dehydromicrosclerodermin D - acetonide in DMSO-d_6

NMR Data – Microsclerodermin L

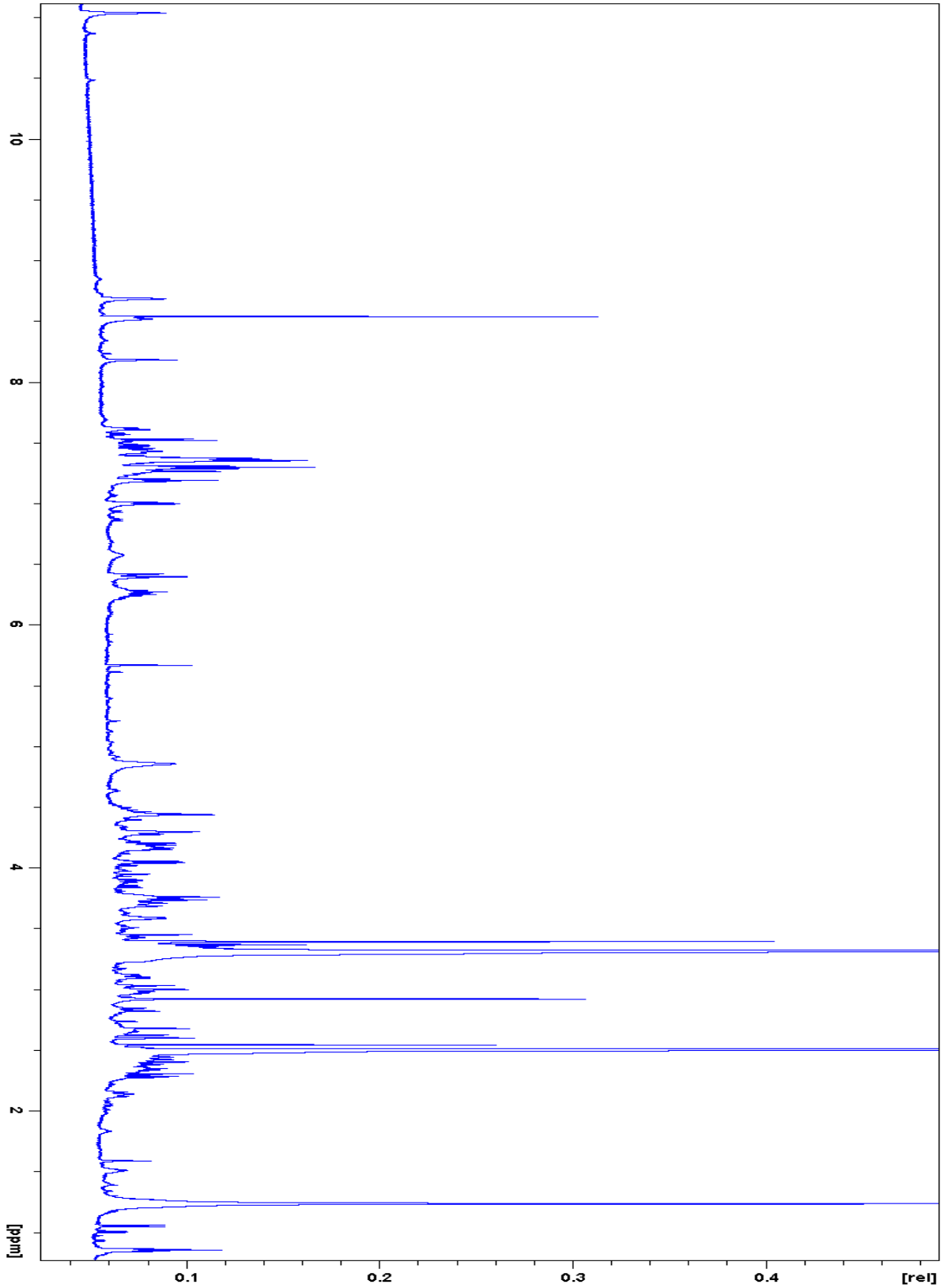


Figure 36: ^{1}H NMR spectrum of microsclerodermin L in DMSO-d_6

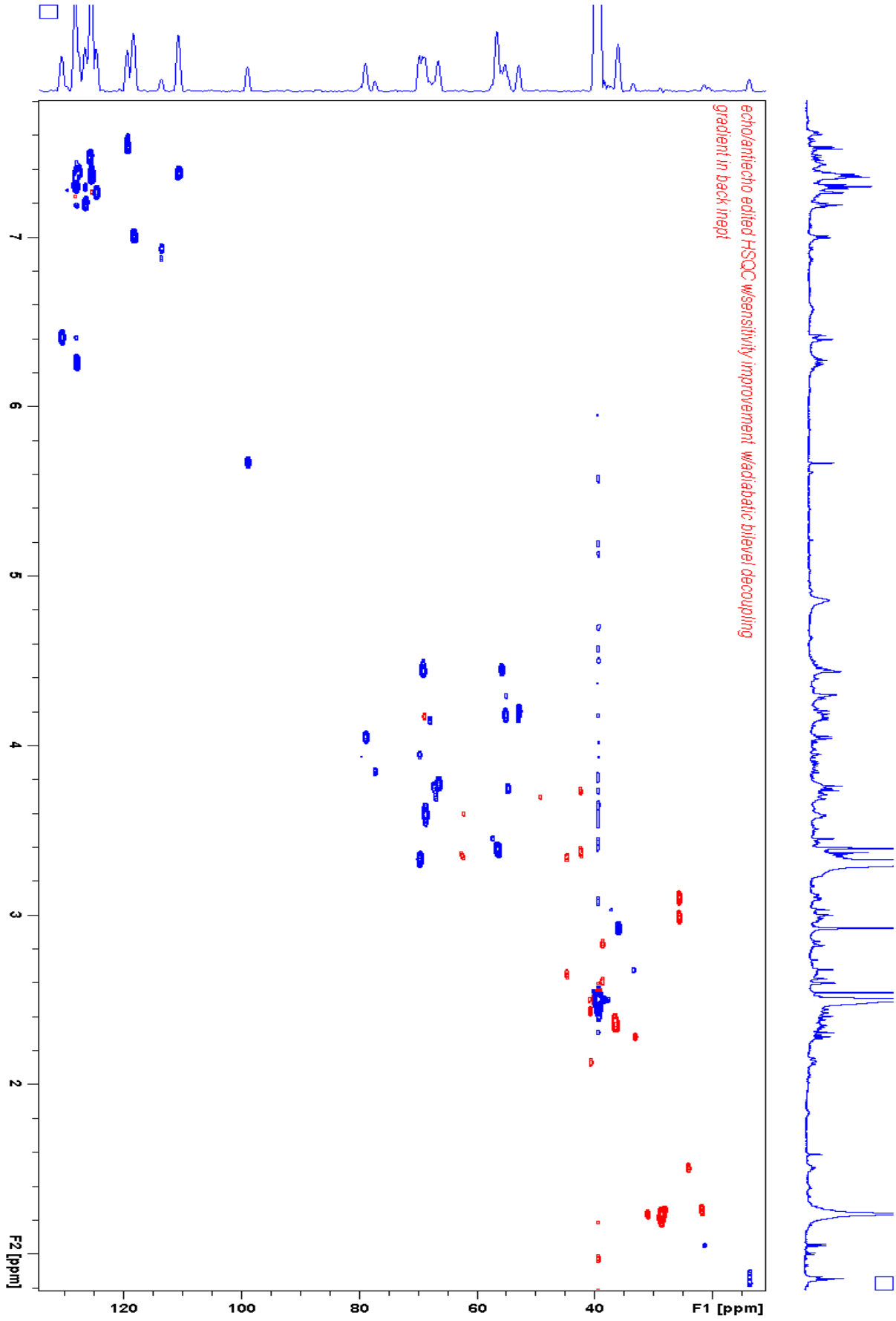


Figure 37: HSQC spectrum of microsclerodermin L in DMSO-d_6

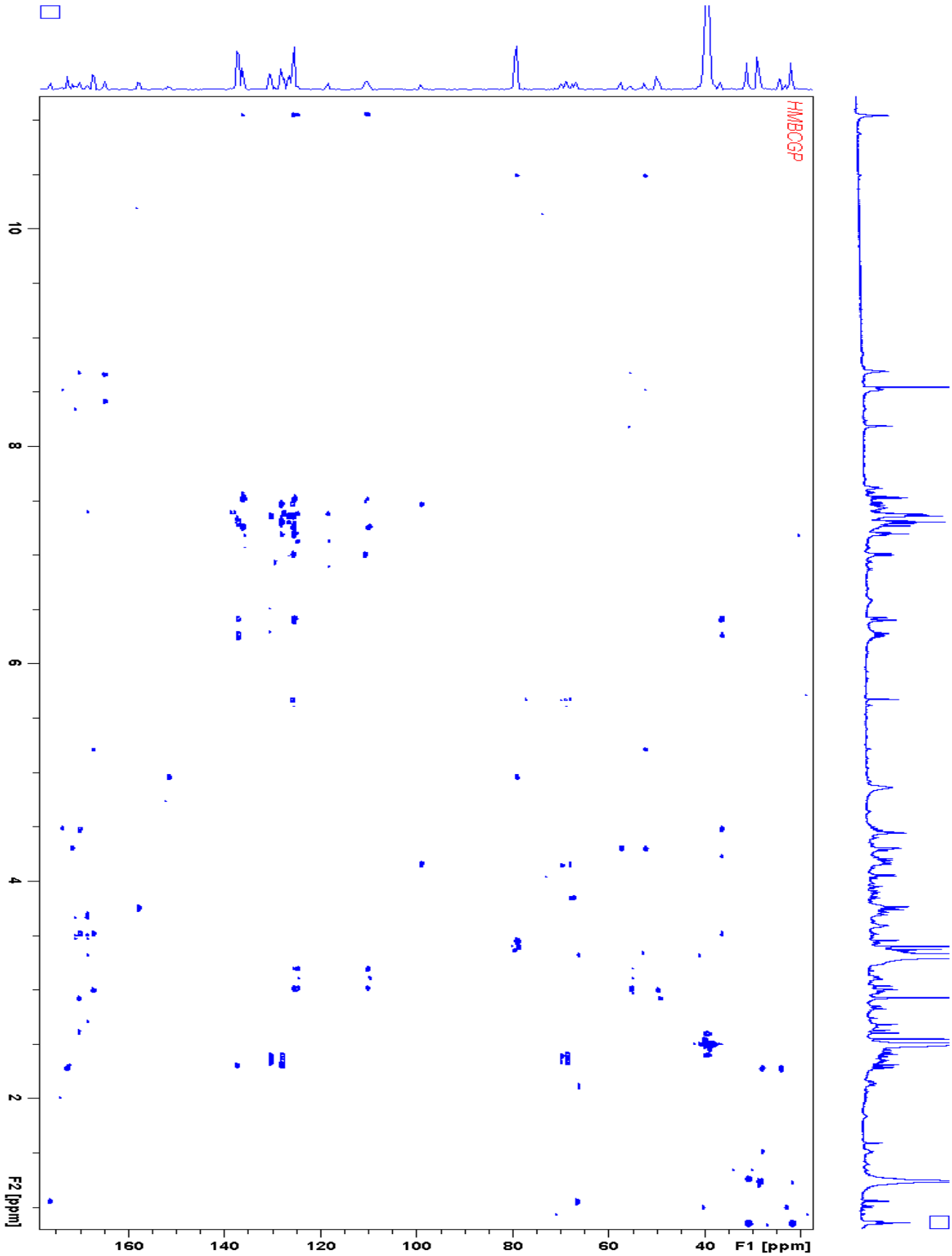


Figure 38: HMBC spectrum of microsclerodermin L in DMSO-d_6

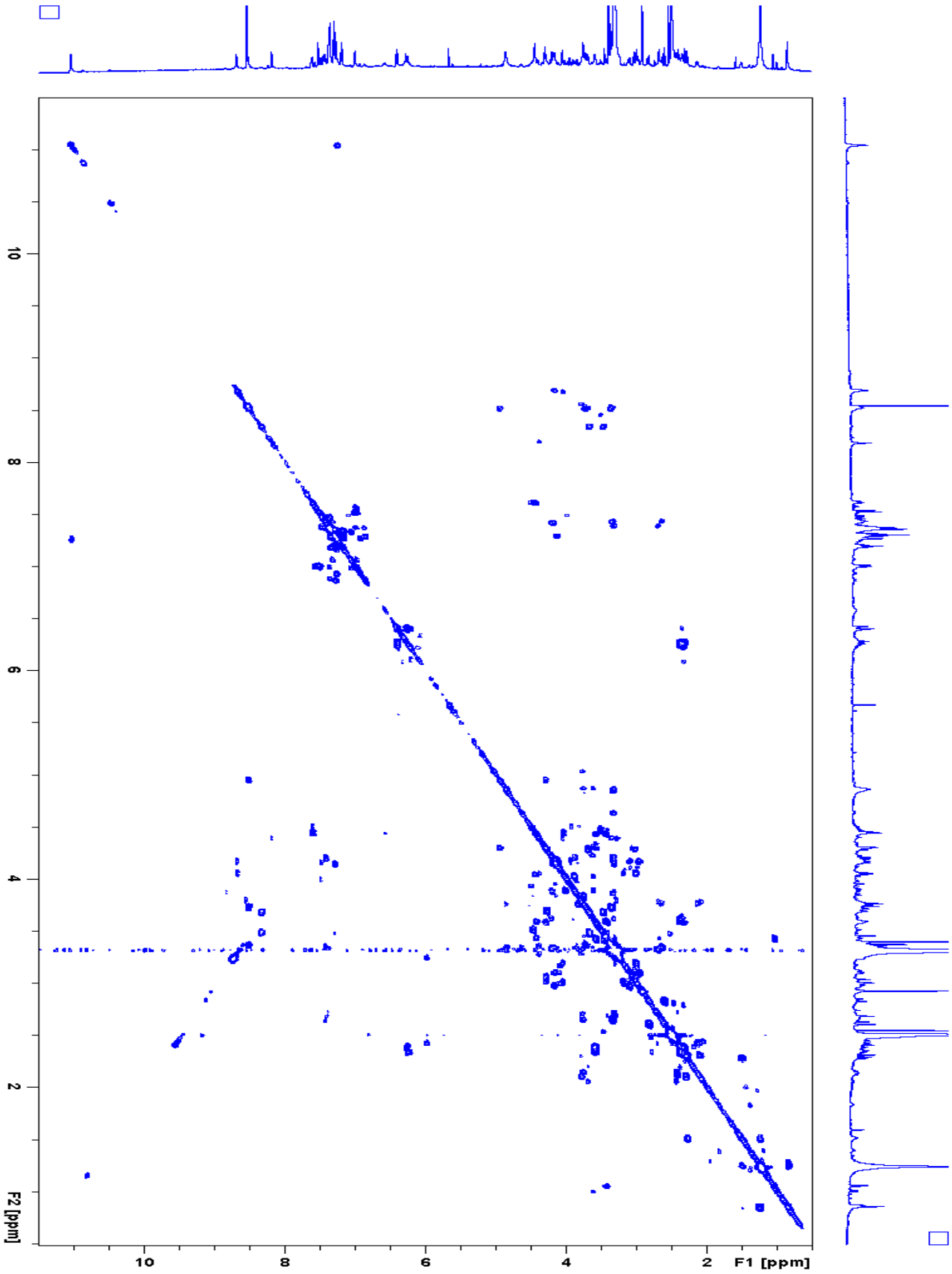


Figure 39: COSY spectrum of microsclerodermin L in DMSO-d_6

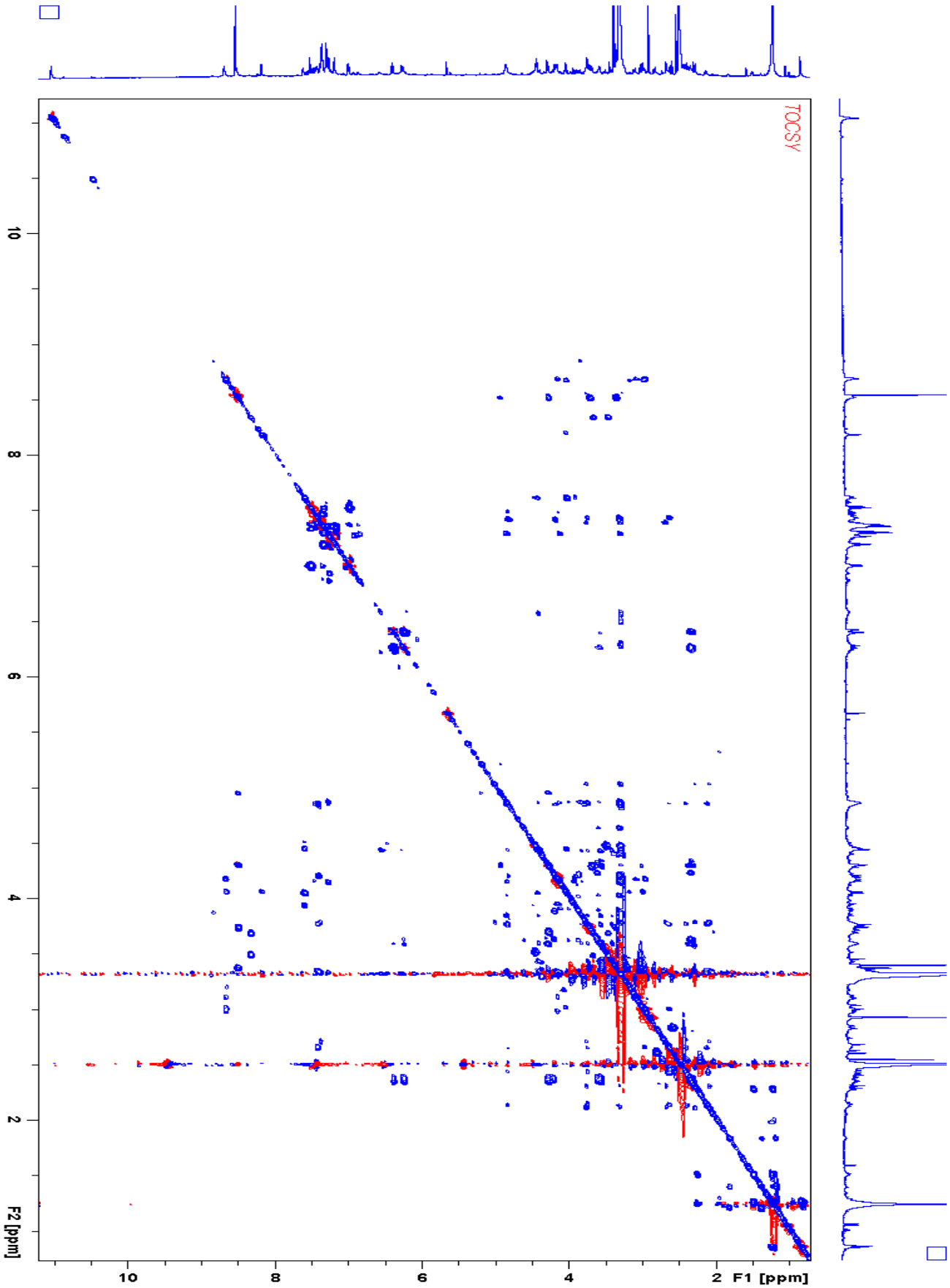


Figure 40: 2D TOCSY spectrum of microsclerodermin L in DMSO-d_6

NMR Data – Microsclerodermin M

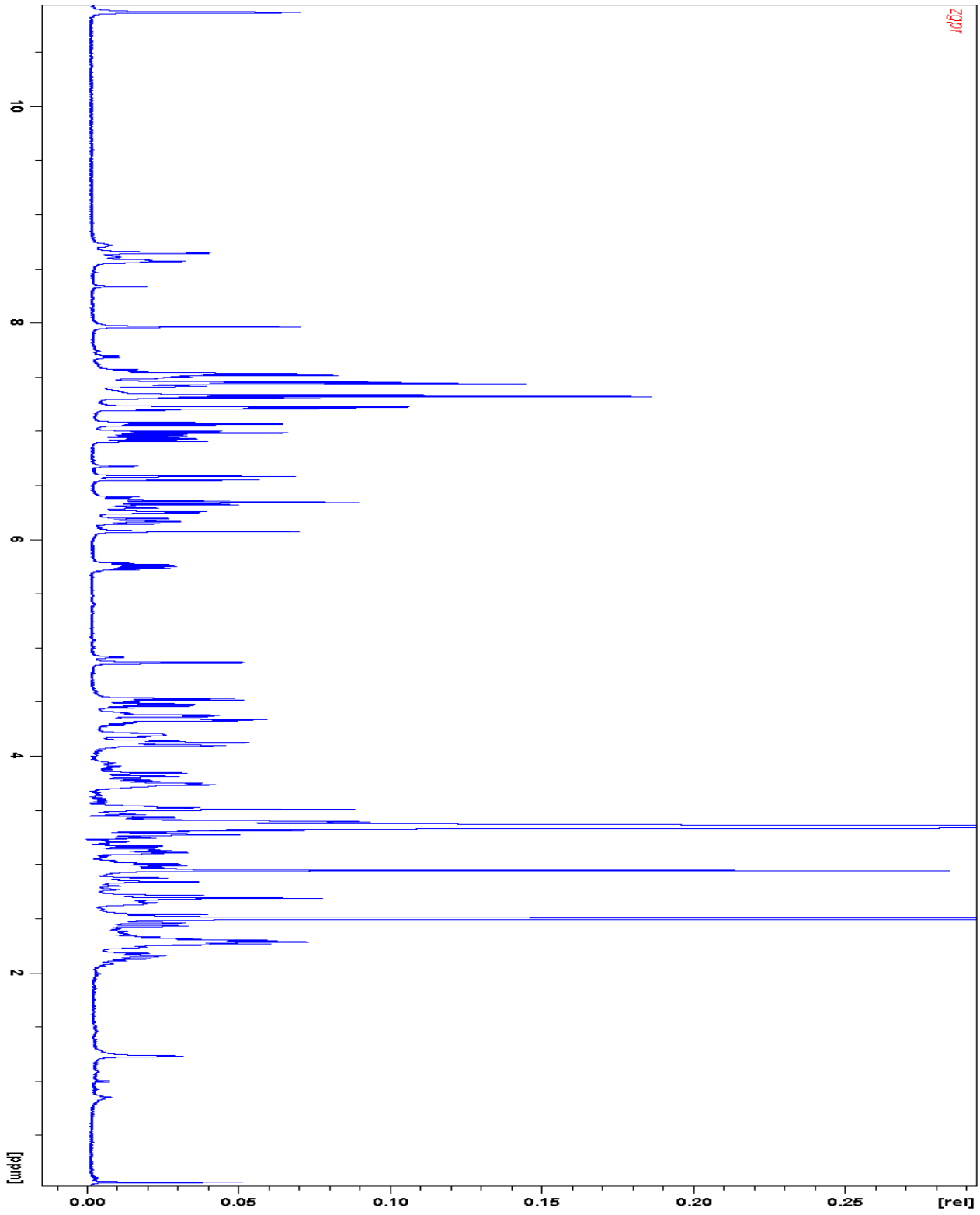


Figure 41: ^{1}H NMR spectrum of microsclerodermin M in DMSO-d_6

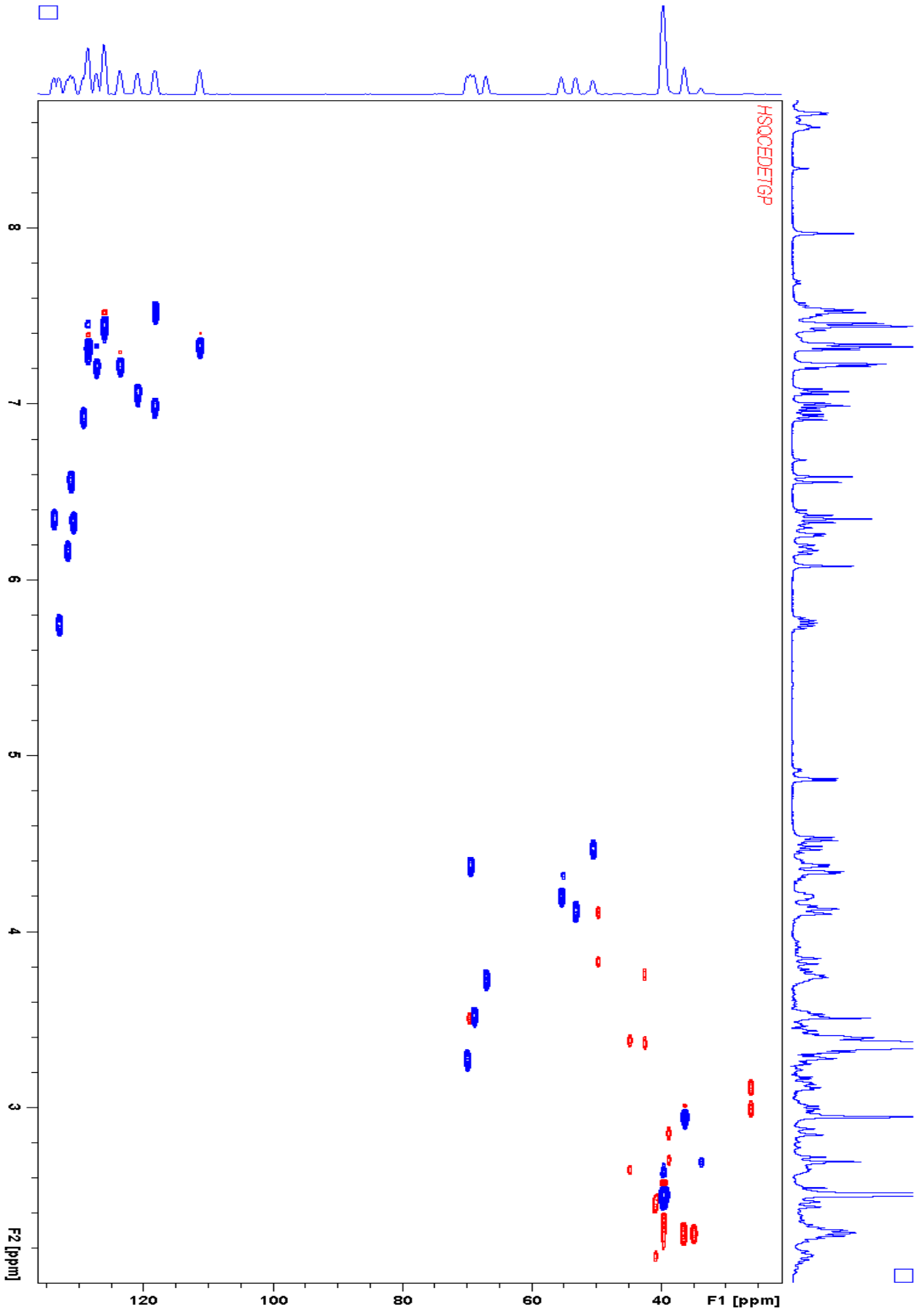


Figure 42: HSQC spectrum of microsclerodermin M in DMSO-d_6

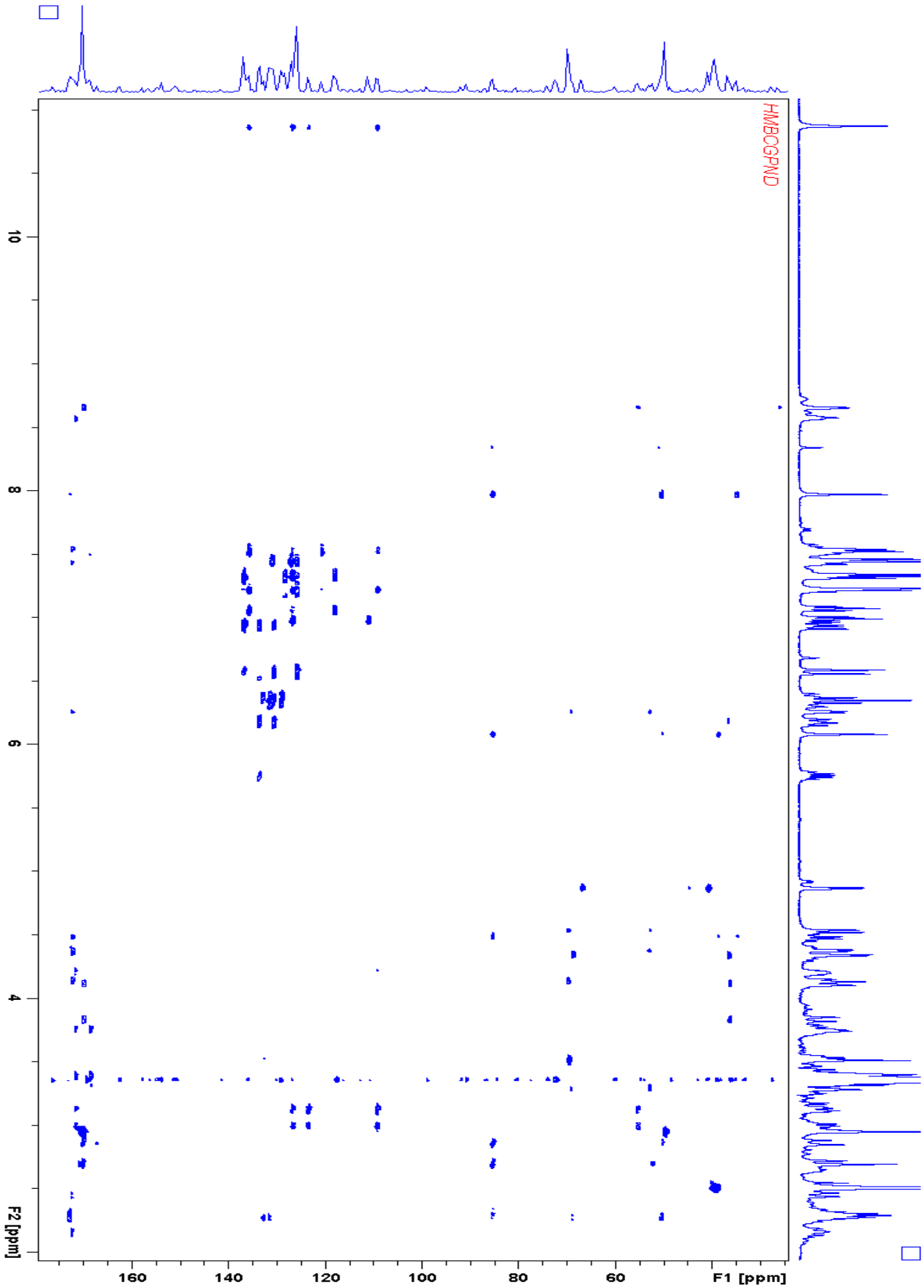


Figure 43: HMBC spectrum of microsclerodermin M in DMSO-d_6

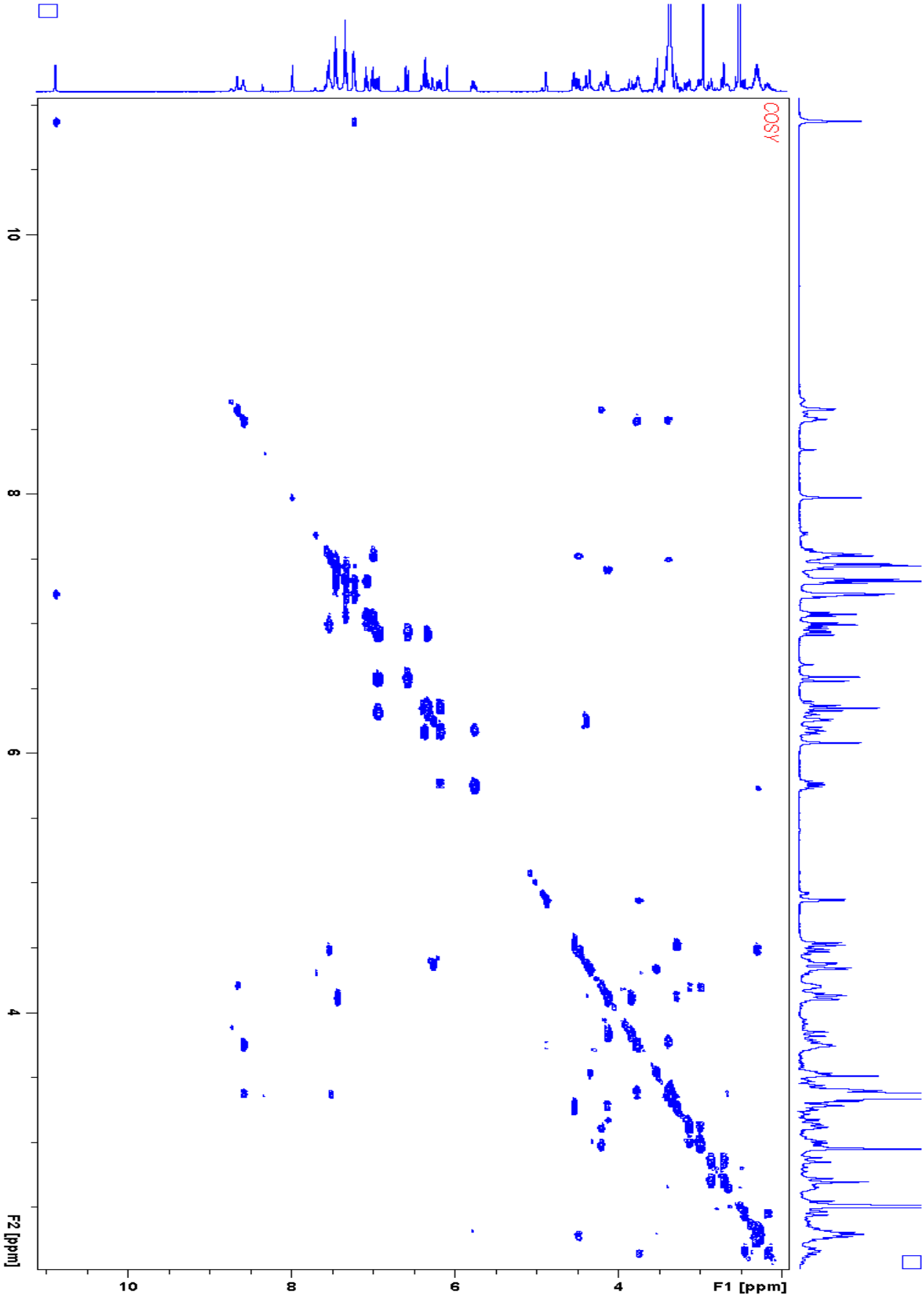


Figure 44: COSY spectrum of microsclerodermin M in DMSO-d_6

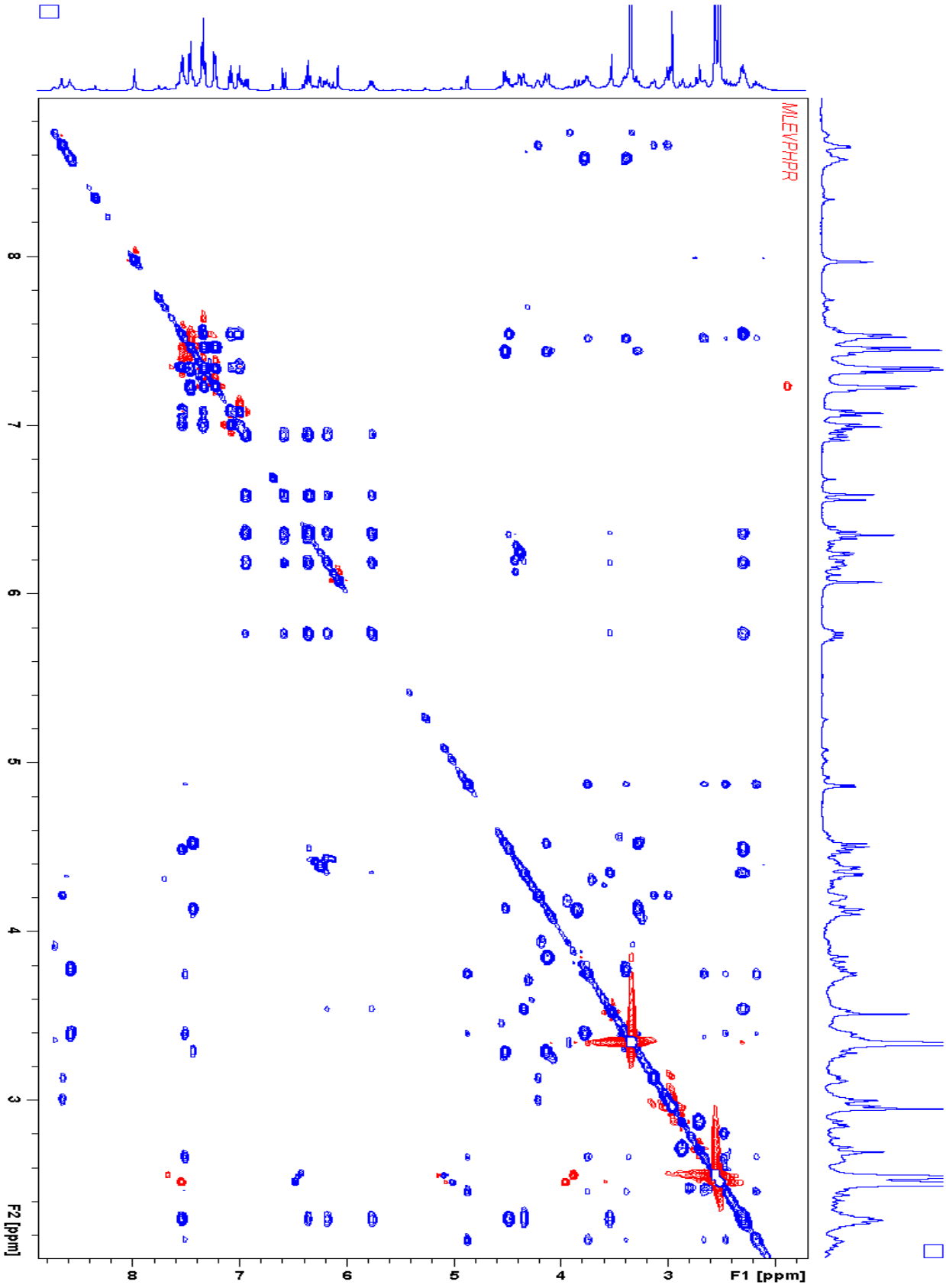


Figure 45: 2D TOCSY spectrum of microsclerodermin M in DMSO-d_6

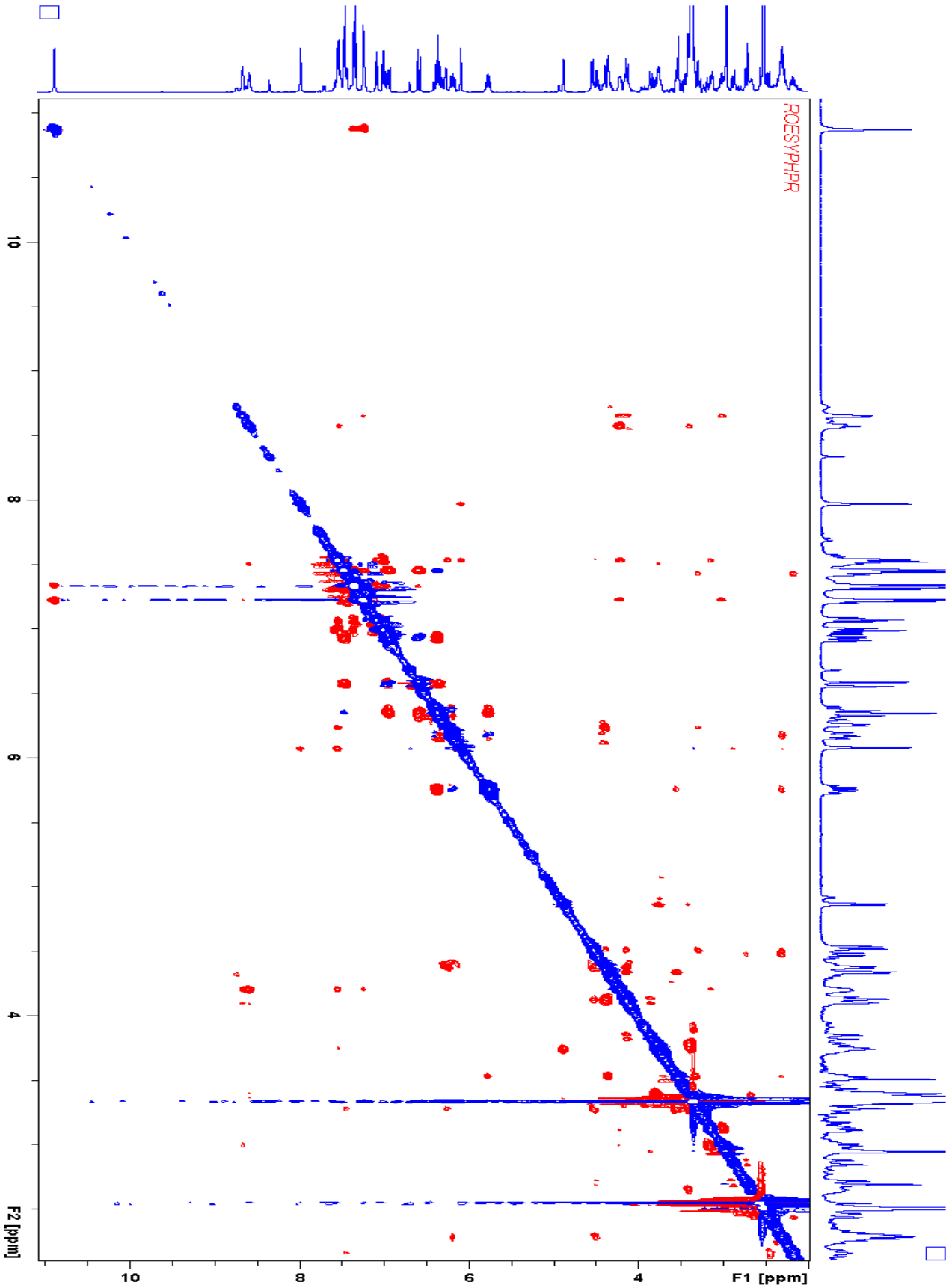


Figure 46: 2D TOCSY spectrum of microsclerodermin M in DMSO-d_6

NMR Data – Dehydromicrosclerodermin M acetonide

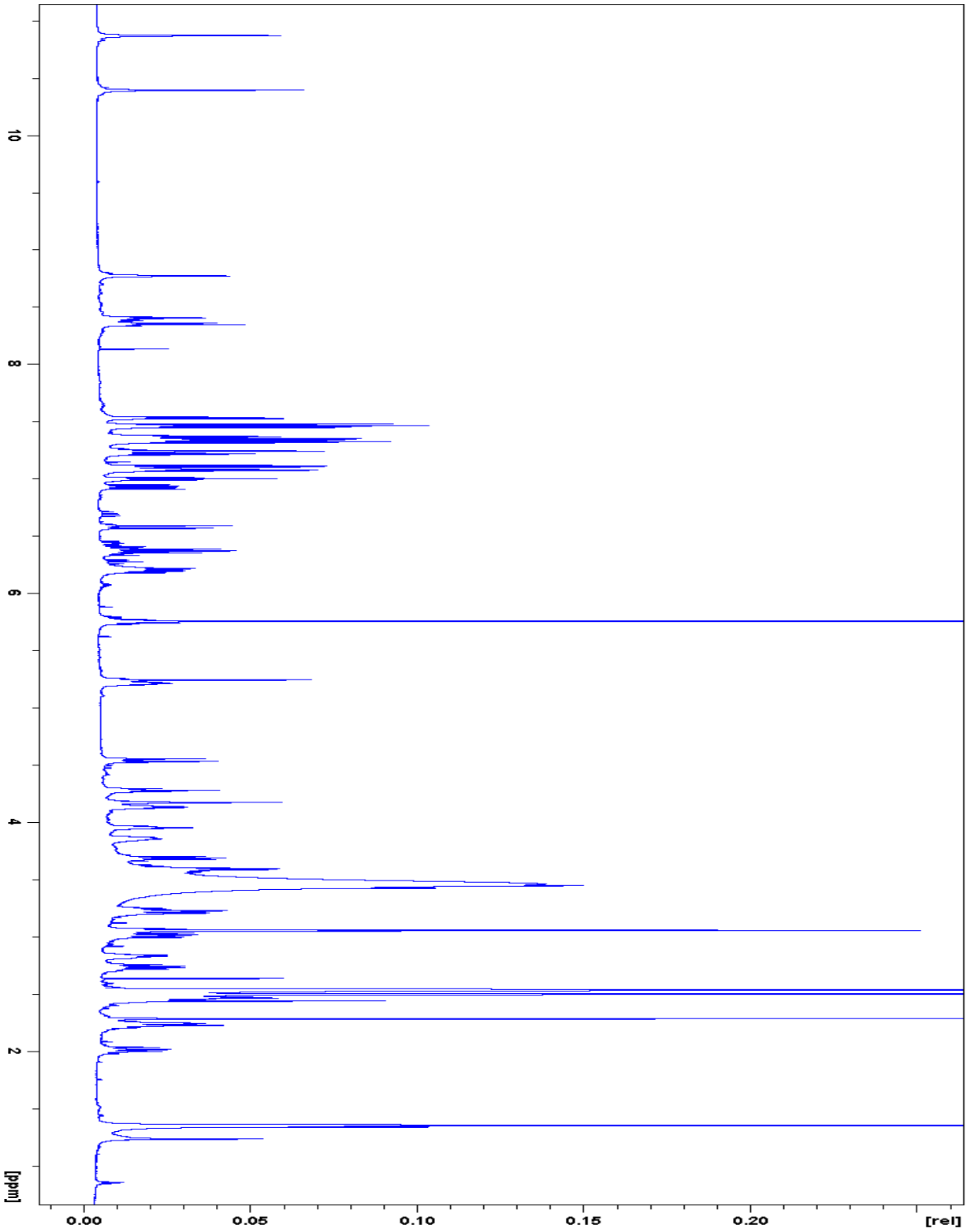


Figure 47: ^{1}H NMR spectrum of acetonide of dehydromicrosclerodermin M in DMSO-d_6

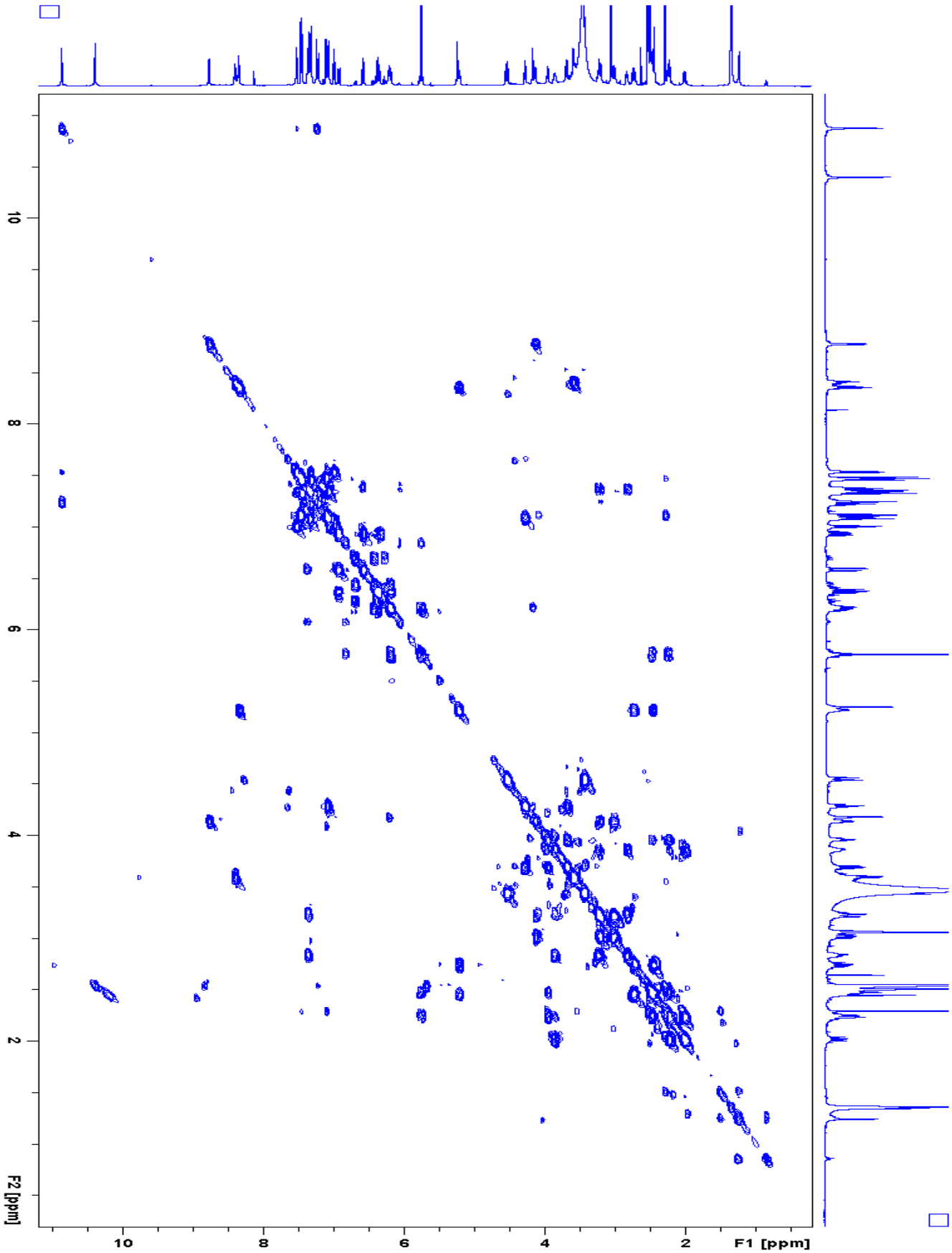


Figure 48: COSY spectrum of acetonide of dehydromicroscl erodermin M in DMSO-d_6

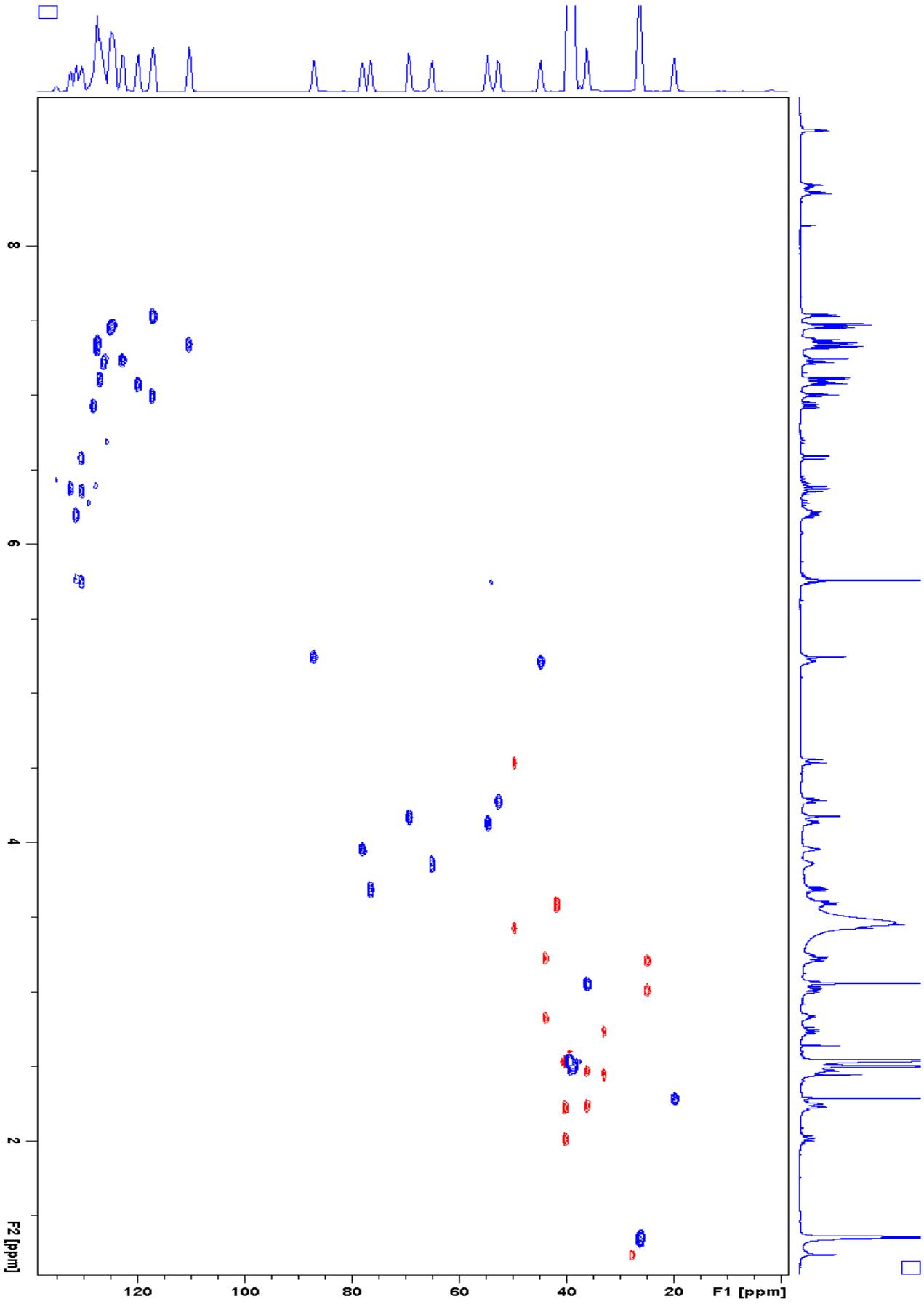


Figure 49: HSQC spectrum of acetonide of dehydromicroscleroderm in M DMSO-d_6